Value-at-Risk bounds with two-sided dependence information

Thibaut Lux [a,1,*], Ludger Rüschendorf [b,2]

ABSTRACT

Value-at-Risk bounds for aggregated risks have been derived in the literature in settings where besides the marginal distributions of the individual risk factors one-sided bounds for the joint distribution respectively the copula of the risks are available. In applications it turns out that these improved standard bounds on Value-at-Risk tend to be too wide to be relevant for practical applications, especially when the number of risk factors is large or when the dependence restriction is not strong enough. In this paper, we develop a method to compute Value-at-Risk bounds when besides the marginal distributions of the risk factors, two-sided dependence information in form of an upper and a lower bound on the copula of the risk factors is available. The method is based on a relaxation of the exact dual bounds which we derive by means of the Monge–Kantorovich transportation duality. In several applications we illustrate that two-sided dependence information typically leads to strongly improved bounds on the Value-at-Risk of aggregations.

KEYWORDS: Value-at-Risk, model uncertainty, copulas, Fréchet-Hoeffding bounds, optimal transport, duality

AUTHORS INFO

[a] *Institute of Financial Mathematics, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium*
[b] *Institute of Mathematics, University of Freiburg, Eckerstraße 1, 79104 Freiburg, Germany*
[1] thibaut.lux@mail.tu-berlin.de
[2] ruschen@stochastik.uni-freiburg.de

[*] *We thank Peter Bank and Ruodu Wang, for useful discussions during the work on these topics. TL would also like to thank Daniel Bartl and Michael Kupper for the instructive conversations in the context of their work on a related topic. TL acknowledges the financial support from the DFG Research Training Group 1845 "Stochastic Analysis with Applications in Biology, Finance and Physics".*

PAPER INFO

AMS CLASSIFICATION: *91B30 (primary), 62E17, 60E15*
JEL CLASSIFICATION: *C02, C63, D80, G31*

1. Introduction

In order to increase the resilience of the financial sector to periods of heightened stress, new regulatory provisions require the computation of robust risk estimates as a key ingredient in the determination of capital reserves; see e.g. Board of Governors of the Federal Reserve System (2011). This in turn calls for new methods to compute risk estimates with partial information about the distribution of the underlying risk drivers. While the computation of portfolio risk estimates from a given model (distribution) for the risk factors poses primarily computational difficulties, fundamentally different challenges arise once we discard the assumption of completely specified model for the risks and consider the framework of model ambiguity. Ambiguity in the Knightian sense occurs whenever uncertainty about the joint law of the risks is introduced. In practice, such uncertainty may stem e.g. from a shortage of historical data to estimate the possibly high-dimensional distribution of the risk factors, or from risk dynamics that are too volatile as to be described adequately by a single model. In this situation, practitioners face the challenge to quantify the portfolio risk in the absence of a completely specified model for the underlying factors. This typically involves the computation of worst-case estimates that correspond to the maximal risk over all possible models that are compatible with reliable information or certain views about the risk factors. These estimates are considered robust with respect to the class of admissible models. In this paper, we present a novel approach to compute robust Value-at-Risk (VaR) estimates for portfolios in the presence of model uncertainty. Thereby, we assume only partial information about the distribution of the risk factors that is available e.g. in the form of reliable estimates or expert views. In applications, we show that our approach provides ways to compute model risk estimates that comply with the five fundamental criteria for robust scenario aggregation presented in Cambou and Filipović (2015), namely (1) no penalty for conservative internal models (2) focus on tail loss (3) control over distance from internal model (4) robustness of capital requirements and (5) tractability.

A significant part of the literature on dependence uncertainty focuses on the marginals-only case, where merely the marginal distributions of $\mathbf{X}$ are known and no information at all about the dependence structure between its constituents is available. In this case, sharp VaR bounds can be obtained by the Rearrangement Algorithm (RA) introduced in Puccetti and Rüschendorf (2012a) and Embrechts, Puccetti, and Rüschendorf (2013). For a presentation of general results in the marginals-only case see Embrechts et al. (2013). The complete absence of information on the dependence structure however leads typically to very wide risk bounds that are not sufficiently informative for practical applications.

This observation has led to a series of papers discussing VaR bounds with additional dependence information in the form of a one-sided, upper or lower bound on the joint distribution function of the risks. The associated VaR bounds are in the literature referred to as *improved standard bounds*. For the ample literature on this we refer to Williamson and Downs (1990), Denuit, Genest, and Marceau (1999), Embrechts, Höing, and Juri (2003), Puccetti and Rüschendorf (2012b), Bignozzi, Puccetti, and Rüschendorf (2015), Bernard and Vanduffel (2015), Puccetti, Rüschendorf, and Manko (2016) and Lux and Papapantoleon (2016). As a result it has been found that this kind of information leads to reasonably narrow risk estimates when the one-sided bounds describe strong enough positive or negative dependence among the risk factors or when the dimension d is relatively small. In higher dimensions the bounds remain often too wide as to be of practical relevance.

To obtain improved VaR bounds, we consider in this paper two-sided constraints on the joint distribution function of the risk vector $\mathbf{X}$. Our main contribution is the development of a method to incorporate two-sided bounds on the copula of $\mathbf{X}$ in order to obtain substantially improved VaR estimates in comparison to the case where only one-sided information is available. For the derivation of two-sided bounds on the copula from partial information about the distribution of the risks we refer to Rachev and Rüschendorf (1994), Nelsen, Quesada-Molina, Rodriguez-Lallena, and Ubeda-Flores (2001), Nelsen (2006, Sec. 3.2.3), Tankov (2011), Lux and Papapantoleon (2015, 2016) and Puccetti et al. (2016).

In the first step we derive an exact dual representation of the VaR bounds over the constrained class of distributions using the Monge-Kantorovich duality theory. The dual bounds however are both analytically and numerically intractable. Nevertheless, the dual formulation allows us to obtain a reduced and tractable optimization scheme for the computation of VaR bounds with two-sided dependence information. Our scheme corresponds to an optimization over a suitable subset of admissible functions for the duals. A similar approach was taken by Embrechts and Puccetti (2006) who derived a numerical procedure from the dual bounds in the marginals-only case, i.e. when only the marginal distributions are known and no information about the dependence structure at all is available. The resulting VaR bounds are no longer sharp in general. Nevertheless, we show that our reduced scheme yields asymptotically sharp risk bounds in the certainty limit, i.e. when uncertainty becomes arbitrarily small. Moreover, we illustrate in numerical examples that our method produces reasonable results also in higher dimensions. In particular, our VaR estimates are significantly tighter than the improved standard bounds, based on one-sided information.

The paper is structured as follows: In Section 2, we develop the relevant notions and give a rigorous definition of the kind of model risk that we address. Section 3 is devoted to the derivation of the dual risk bounds and the proof of strong duality. In Section 4, we then present numerical schemes for the upper and the lower VaR bounds based on the dual form of the risk estimates. Moreover, we prove asymptotic sharpness of the bounds as uncertainty vanishes. We conclude, in Section 5, with a graphical illustration of our numerical scheme as well as several examples highlighting the performance of the improved VaR bounds.

2. Bounds on Value-at-Risk using copula information

In this paper we consider an $\mathbb{R}^d$-valued random vector of risks $\mathbf{X} = (X_1, \dots, X_d)$ and an aggregation function $\psi\colon \mathbb{R}^d \to \mathbb{R}$. We want to compute the VaR of the aggregation $\psi(\mathbf{X})$. The VaR of $\psi(\mathbf{X})$ relates to the quantile function in the following way: when $\psi(\mathbf{X}) \sim F_\psi$ then the VaR of $\psi(\mathbf{X})$ for a certain confidence level $\alpha \in (0,1)$ is given by the quantity[1]

$$\mathrm{VaR}_\alpha(\psi(\mathbf{X})) = F_\psi^{-1}(\alpha) = \inf\{x \in \mathbb{R}\colon F_\psi(x) > \alpha\}.$$

Typical levels of α are close to 1, assuming that risks (or losses) correspond to the right tail of the distribution. The most commonly considered aggregation function ψ is the sum of the individual risks $X_1 + \cdots + X_d$, but also the maximum and minimum of the risks, $\max\{X_1, \dots, X_d\}$ and $\min\{X_1, \dots, X_d\}$ are of interest.

We are concerned with the situation of model ambiguity and assume that only partial information about the distribution of $\mathbf{X}$ is available. Frequently, the univariate distributions of the constituents $X_1, ..., X_d$ are known or can be estimated while the dependence structure between the individual components is at best partially known. This form of model ambiguity is referred to as *dependence uncertainty*. It is assumed that the unknown joint distribution F of $\mathbf{X}$ is in the Fréchet class $\mathcal{F}(F_1, ..., F_d)$ of d-dimensional distribution functions with marginals $F_1, ..., F_d$. Then it follows from Sklar's Theorem that F can be expressed as

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \quad \text{for all } (x_1, ..., x_d) \in \mathbb{R}^d, \tag{2.1}$$

[1] Sometimes, VaR is defined using '$\geq$' instead of '$>$' for the definition of the generalized inverse. Our formulation here guarantees suitable continuity properties of objective functions used in the following sections. For a detailed discussion of the consequences of either VaR formulation c.f. Embrechts and Hofert (2013).

for a copula C. This implies that dependence uncertainty is in fact uncertainty about the copula of $\mathbf{X}$.

When no information about the risk vector besides its marginal distributions is provided, then every copula yields a possible joint distribution for $\mathbf{X}$ via (2.1). In this situation, VaR bounds for aggregations $\psi(\mathbf{X})$ over the set of all copulas are typically too wide to be relevant in practice. In addition, it is rarely the case that no information at all about the dependence structure of the risk vector is available, since partial information such as e.g. Kendall's tau or correlations between the risk factors can be estimated or inferred with sufficient accuracy. Such additional information can be translated into a lower and an upper bound on the copula of $\mathbf{X}$; see e.g. Rachev and Rüschendorf (1994), Nelsen (2006) and Tankov (2011) for $d = 2$ or Lux and Papapantoleon (2015, 2016) and Puccetti et al. (2016) for $d > 2$. Our approach will allow us to translate these improved Fréchet–Hoeffding bounds into bounds on VaR for the aggregation $\psi(\mathbf{X})$. In fact, we consider first a more general problem. Instead of bounding VaR we derive bounds on the expectation of a general functional $\varphi(\mathbf{X})$ when bounds on the copula of $\mathbf{X}$ are available.

In the following let the marginal distributions of the risk vector $(X_1, ..., X_d)$ be fixed and denoted by $F_1, ..., F_d$. With this specification, we denote the expectation operator for a measurable $\varphi\colon \mathbb{R}^d \to \mathbb{R}$ and an $\mathbb{R}^d$-valued random vector $\mathbf{X}$ with copula C, by

$$\mathbb{E}_C[\varphi] = \int_{\mathbb{R}^d} \varphi(x_1, ..., x_d) \, \mathrm{d}C(F_1(x_1), ..., F_d(x_d)).$$

Using $C \leq C'$ to refer to the pointwise inequality between d-variate functions, we define the *generalized Fréchet functionals* with two-sided constraints by

$$\underline{P}_\varphi := \inf\left\{\mathbb{E}_C[\varphi] \colon C \in \mathcal{C}^d, \underline{Q} \leq C \leq \overline{Q}\right\}, \tag{2.2}$$

$$\overline{P}_\varphi := \sup\left\{\mathbb{E}_C[\varphi] \colon C \in \mathcal{C}^d, \underline{Q} \leq C \leq \overline{Q}\right\}, \tag{2.3}$$

where $\mathcal{C}^d$ is the set of all d-copulas and $\overline{Q}, \underline{Q}$ are quasi-copulas with $\underline{Q} \leq \overline{Q}$. The generalized Fréchet functionals describe the maximal respectively minimal influence of dependence on the expectation of φ in the Fréchet class with two-sided constraints. The notion of quasi-copulas generalizes the copula concept as follows:

Definition 2.1. A function $Q\colon [0,1]^d \to [0,1]$ is a *d-quasi-copula* if the following properties hold:

$(\mathbf{QC1})$ Q satisfies, for all $i \in \{1, \dots, d\}$, the boundary conditions

$$Q(u_1, \dots, u_i = 0, \dots, u_d) = 0 \quad \text{and} \quad Q(1, \dots, 1, u_i, 1, \dots, 1) = u_i.$$

$(\mathbf{QC2})$ Q is increasing in each argument.

$(\mathbf{QC3})$ Q is Lipschitz continuous, i.e. for all $\mathbf{u}, \mathbf{v} \in [0,1]^d$

$$|Q(u_1, \dots, u_d) - Q(v_1, \dots, v_d)| \leq \sum_{i=1}^d |u_i - v_i|.$$

In particular, by allowing quasi-copulas as bounds in the formulation of the generalized Fréchet functionals, we are able to include the lower Fréchet–Hoeffding bound or improved Fréchet–Hoeffding bounds which are quasi-copulas but often fail to be proper copulas. Allowing quasi-copulas as bounds, there might however not exist a copula that complies with the constraints and so $\mathcal{C}_b := \{C \in \mathcal{C}^d : \underline{Q} \le C \le \overline{Q}\} = \emptyset$. Consider e.g. $\overline{Q} = \underline{Q} = W_d$ where W_d is the d-dimensional lower Fréchet–Hoeffding bound, then $\mathcal{C}_b$ is empty whenever $d > 2$. In this case, we set $\overline{P}_\varphi = \infty$ and $\underline{P}_\varphi = -\infty$.

Remark 2.2. When $\underline{Q}$ and $\overline{Q}$ are equal to the lower and upper Fréchet–Hoeffding bound respectively, i.e.

$$\begin{aligned} \underline{Q}(u_1, ..., u_d) &= \max\left\{0, \sum_{i=1}^d u_i - d + 1\right\} =: W_d(\mathbf{u}) \quad \text{and} \\ \overline{Q}(u_1, ..., u_d) &= \min(u_1, ..., u_d) =: M_d \end{aligned} \tag{2.4}$$

then the optimization corresponds to a standard Fréchet problem where only the marginals are known and no information about the dependence structure at all is available. ♦

Remark 2.3. Let us point out that the Lipschitz property (**QC3**) of the bounds is not a limiting constraint in applications. Bounds on copulas derived from partial information, that are known in the literature, are essentially pointwise infima or suprema of sets of copulas, i.e. $\underline{Q}(\mathbf{u}) = \inf\{C(\mathbf{u}) : C \in \mathcal{C}\}$, for $\mathcal{C}$ being some constrained set of copulas. Hence, they are in particular Lipschitz continuous. ♦

In order to compute or approximate the bounds $\overline{P}_\varphi$ and $\underline{P}_\varphi$ we proceed as follows: First, we derive a dual representation of the generalized Fréchet functionals yielding sharp bounds on the expectation of $\varphi(\mathbf{X})$ under rather general assumptions on the function φ. Based on the dual representation we then develop a tractable optimization scheme to compute bounds on the expectation for specific functions φ. In particular, the scheme allows us to determine robust VaR estimates of aggregations using copula bounds.

3. Dual representation of generalized Fréchet functionals with two-sided bounds on the copula

In this section we establish a dual characterization of the generalized Fréchet functionals $\overline{P}_\varphi$ and $\underline{P}_\varphi$ and prove strong duality between the two formulations. To this end, we introduce the class

$$\mathcal{R} := \left\{h = \sum_{n=1}^k \alpha_n \Lambda_{\mathbf{u}^n} : k \in \mathbb{N}, \alpha_1, ..., \alpha_k \ge 0, \mathbf{u}^1, ..., \mathbf{u}^k \in \overline{\mathbb{R}}^d\right\},$$

where the functions $\Lambda_{\mathbf{u}}$ are of the form

$$\Lambda_{\mathbf{u}} : \mathbb{R}^d \ni (x_1, ..., x_d) \mapsto \mathbb{1}_{x_1 \le u_1, ..., x_d \le u_d}.$$

Here, we convene that the superscript n in $\mathbf{u}^n$ is the index of the sequence $\mathbf{u}^1, ..., \mathbf{u}^k$ and does not refer to the exponentiation of $\mathbf{u}$.

The elements in $\mathcal{R}$ are hence positive, linear combinations of indicator functions of rectangles of the form $(-\infty, u_1] \times \cdots \times (-\infty, u_d]$. Analogously, we denote the lower semicontinuous version of $\Lambda_{\mathbf{u}}$ by

$$\Lambda_{\mathbf{u}}^- \colon (x_1, ..., x_d) \mapsto \mathbb{1}_{x_1 < u_1, ..., x_d < u_d},$$

and for $h = \sum_{n=1}^k \alpha_n \Lambda_{\mathbf{u}^n} \in \mathcal{R}$ we define $h^- := \sum_{n=1}^k \alpha_n \Lambda_{\mathbf{u}^n}^-$.

Note that for a copula C it holds that

$$\mathbb{E}_C[\Lambda_{\mathbf{u}}] = \int_{\mathbb{R}^d} \Lambda_{\mathbf{u}}(x_1, ..., x_d) \, \mathrm{d}C\big(F_1(x_1), ..., F_d(x_d)\big) = C\big(F_1(u_1), ..., F_d(u_d)\big), \tag{3.1}$$

and analogously we obtain that $\mathbb{E}_C[\Lambda_{\mathbf{u}}^-] = C(F_1^-(u_1), ..., F_d^-(u_d))$, where F_i^- is the left-continuous version of F_i for $i = 1, ..., d$.

Moreover, we define, for a quasi-copula Q and $h = \sum_{n=1}^k \alpha_n \Lambda_{\mathbf{u}^n} \in \mathcal{R}$,

$$Q(h) := \sum_{n=1}^k \alpha_n Q\big(F_1(u_1^n), ..., F_d(u_d^n)\big); \quad Q(h^-) := \sum_{n=1}^k \alpha_n Q\big(F_1^-(u_1^n), ..., F_d^-(u_d^n)\big).$$

If $Q = C$ for a copula C, we have that $Q(h) = \mathbb{E}_Q[h]$ as well as $Q(h^-) = \mathbb{E}_Q[h^-]$.

We then guess a dual form of the generalized Fréchet functional $\underline{P}_\varphi$ and prove in Theorem 3.5 below that it is in fact an alternative formulation of $\underline{P}_\varphi$, i.e. there is no duality gap. The intuition behind the dual characterization is that we seek to maximize the expectation over all objective functions which are dominated by φ and whose expectation can be computed with certainty from available information. This leads us to the following dual form of the generalized Fréchet functional $\underline{P}_\varphi$:

$$\begin{aligned} \underline{D}_\varphi = \sup \Big\{ \underline{Q}(h) - \overline{Q}(g^-) + \sum_{i=1}^d \mathbb{E}_i[f_i] \colon & f_i \in \mathcal{L}(F_i), i = 1, ..., d; \\ & h, g \in \mathcal{R} \text{ s.t. } h - g^- + \sum_{i=1}^d f_i \le \varphi \Big\}, \end{aligned} \tag{3.2}$$

where $\mathbb{E}_i[f_i] = \int f_i \, \mathrm{d}F_i$ and $\mathcal{L}(F_i)$ is the class of F_i-integrable functions f_i, i.e. $\mathbb{E}_i[|f_i|] < \infty$, for $i = 1, ..., d$. Analogously, for the upper bound $\overline{P}_\varphi$, the corresponding dual is given by

$$\begin{aligned} \overline{D}_\varphi = \inf \Big\{ \overline{Q}(h^-) - \underline{Q}(g) + \sum_{i=1}^d \mathbb{E}_i[f_i] \colon & f_i \in \mathcal{L}(F_i), i = 1, ..., d; \\ & h, g \in \mathcal{R} \text{ s.t. } h^- - g + \sum_{i=1}^d f_i \ge \varphi \Big\}. \end{aligned} \tag{3.3}$$

Note, that the roles of $\underline{Q}$ and $\overline{Q}$ are reversed in $\underline{D}_\varphi$, i.e. we subtract the sum w.r.t. $\overline{Q}$ from the sum w.r.t. $\underline{Q}$ in the formulation of $\underline{D}_\varphi$ and vice versa for $\overline{D}_\varphi$. In the remainder of this section, we show that strong duality between the generalized Fréchet functionals and the dual characterization holds under mild assumptions on the function φ, so that:

$$\underline{P}_\varphi = \underline{D}_\varphi \quad \text{and} \quad \overline{P}_\varphi = \overline{D}_\varphi.$$

Several approaches to proving duality results of this type have been established in the literature. Rüschendorf (1981) and Gaffke and Rüschendorf (1981) establish duality results for functionals of multivariate random variables with given marginals using a Hahn-Banach separation argument. This method was also extended to some cases with additional or relaxed constraints. A more general duality result, based on a different method, was given in Kellerer (1984). A duality result for the martingale optimal transport problem was established by Beiglböck, Henry-Labordère, and Penkner (2013), using the Kantorovich Duality Theorem combined with a minimax argument. Bartl, Cheredito, Kupper, and Tangpi (2015) derive a general duality result for convex functionals with countably many marginal constraints using the Daniell-Stone Theorem. An account of the history of the Monge-Kantorovich duality theory and associated references can be found in the survey by Rüschendorf (2007) or in the book by Villani (2009).

The proof of our duality theorem 3.2 for the generalized Fréchet functionals is based on the following copula-version of the Kantorovich duality: Let $\varphi\colon \mathbb{R}^d \to \mathbb{R}$ be lower semicontinuous and such that for some $g_i \in \mathcal{L}(F_i)$, $i = 1, ..., d$ we have

$$\sum_{i=1}^{d} g_i(x_i) \geq |\varphi(x_1, ..., x_d)| \quad \text{for all } (x_1, ..., x_d) \in \mathbb{R}^d,$$

then it holds that

$$\inf \left\{ \mathbb{E}_C[\varphi]\colon C \in \mathcal{C}^d \right\} = \sup \left\{ \sum_{i=1}^{d} \mathbb{E}_i[f_i]\colon f_i \in \mathcal{L}(F_i),\ i = 1, ..., d,\ \sum_{i=1}^{d} f_i \leq \varphi \right\}. \tag{3.4}$$

Equation (3.4) follows immediately from the duality result in Villani (2009, Ch. 5). We remark that more general versions of the Kantorovich duality, e.g. for φ being merely a Baire-function or product-measurable, exist in the literature; see e.g. Rachev and Rüschendorf (1998, Ch. 2).

Moreover, we will make use of the classical Minimax Theorem of Ky-Fan.

Lemma 3.1 (Minimax Theorem). *Let B_1 be a compact convex subset of a topological vector space V_1 and B_2 be a convex subset of a vector space V_2. If $f\colon B_1 \times B_2 \to \mathbb{R}$ is such that*

1. *$f(\cdot, b_2)$ is lower semicontinuous and convex on B_1 for all $b_2 \in B_2$,*
2. *$f(b_1, \cdot)$ is concave on B_2 for all $b_1 \in B_1$,*

then

$$\inf_{b_1 \in B_1} \sup_{b_2 \in B_2} f(b_1, b_2) = \sup_{b_2 \in B_2} \inf_{b_1 \in B_1} f(b_1, b_2).$$

With these results, we are now in the position to establish our main duality theorem.

Theorem 3.2 (Dual bounds with two-sided dependence information). *Let* $\varphi\colon \mathbb{R}^d \to \mathbb{R}$ *be such that*

$$\sum_{i=1}^{d} g_i(x_i) \geq |\varphi(x_1, ..., x_d)| \quad \textit{for all } (x_1, ..., x_d) \in \mathbb{R}^d, \tag{3.5}$$

for some elements $g_i \in \mathcal{L}(F_i)$*,* $i = 1, ..., d$*. Moreover, assume that there exists a copula* $C \in \mathcal{C}^d$ *with* $\underline{Q} \leq C \leq \overline{Q}$*. Then if* φ *is lower semicontinuous the following duality holds:*

$$\underline{P}_\varphi = \underline{D}_\varphi.$$

When φ *is upper semicontinuous the following duality holds:*

$$\overline{P}_\varphi = \overline{D}_\varphi.$$

Moreover, there exist copulas $\underline{C}, \overline{C}$ *such that* $\mathbb{E}_{\underline{C}}[\varphi] = \underline{P}_\varphi$ *and* $\mathbb{E}_{\overline{C}}[\varphi] = \overline{P}_\varphi$*.*

Proof. We show that the statement holds for the lower bound, i.e. $\underline{D}_\varphi = \underline{P}_\varphi$. The proof for the upper bound can be derived by applying analogous arguments to the function $-\varphi$.

First, assume that φ is bounded and continuous. By f_i we refer to functions in $\mathcal{L}(F_i)$. It follows that

$$\underline{D}_\varphi = \sup_{h,g \in \mathcal{R}} \sup_{\substack{f_1,...,f_d \\ h - g^- + \sum_{i=1}^d f_i \leq \varphi}} \left\{ \underline{Q}(h) - \overline{Q}(g^-) + \sum_{i=1}^{d} \mathbb{E}_i[f_i] \right\} \tag{3.6}$$

$$= \sup_{h,g \in \mathcal{R}} \sup_{\substack{f_1,...,f_d \\ \sum_{i=1}^d f_i \leq \varphi - h + g^-}} \left\{ \underline{Q}(h) - \overline{Q}(g^-) + \sum_{i=1}^{d} \mathbb{E}_i[f_i] \right\} \tag{3.7}$$

$$= \sup_{h,g \in \mathcal{R}} \inf_{C \in \mathcal{C}^d} \left\{ \underline{Q}(h) - \overline{Q}(g^-) + \mathbb{E}_C[\varphi - h + g^-] \right\} \tag{3.8}$$

$$= \sup_{h,g \in \mathcal{R}} \inf_{C \in \mathcal{C}^d} \left\{ \big(\underline{Q}(h) - C(h)\big) - \big(\overline{Q}(g^-) - C(g^-)\big) + \mathbb{E}_C[\varphi] \right\} \tag{3.9}$$

$$= \inf_{C \in \mathcal{C}^d} \left\{ \sup_{h,g \in \mathcal{R}} \left\{ \big(\underline{Q}(h) - C(h)\big) - \big(\overline{Q}(g^-) - C(g^-)\big) \right\} + \mathbb{E}_C[\varphi] \right\} \tag{3.10}$$

$$= \inf_{\substack{C \in \mathcal{C}^d \\ \underline{Q}(h) \leq C(h) \leq \overline{Q}(h),\ \forall h \in \mathcal{R}}} \mathbb{E}_C[\varphi] \tag{3.11}$$

$$= \inf_{\underline{Q} \leq C \leq \overline{Q}} \mathbb{E}_C[\varphi] = \underline{P}_\varphi. \tag{3.12}$$

Equation (3.8) follows from an application of the Kantorovich Duality Theorem to the function $\varphi' := \varphi - h + g^-$; see equation (3.4). Note, that the application of the theorem is justified since

φ' is lower semicontinuous being the sum of the lower semicontinuous functions φ, $-h$ and g^-. Moreover, since h and g are of the form

$$h = \sum_{n=1}^{k} \alpha_n \Lambda_{\mathbf{u}^n}, \quad g = \sum_{n=1}^{m} \beta_n \Lambda_{\mathbf{v}^n}$$

for $\alpha_1, ..., \alpha_k, \beta_1, ..., \beta_m \in \mathbb{R}_+$, we obtain

$$|(\varphi + h - g^-)(x_1, ..., x_d)| \leq \sum_{i=1}^{d} g_i(x_i) + \sum_{n=1}^{k} \alpha_n + \sum_{n=1}^{m} \beta_n.$$

Equation (3.9) then follows by rearranging the terms, using the linearity of the expectation and the definition of the operator $C(h)$ for $h \in \mathcal{R}$. Now, applying the Minimax Theorem 3.1 to the function

$$f \colon \mathcal{C}^d \times \mathcal{R}^2 \ni \big(C, (h, g)\big) \mapsto \big(\underline{Q}(h) - C(h)\big) - \big(\overline{Q}(g^-) - C(g^-)\big) + \mathbb{E}_C[\varphi]$$

yields equation (3.10). Note, that the requirements of Theorem 3.1 are satisfied, since

$$\mathcal{C}_b = \big\{C \in \mathcal{C}^d \colon \underline{Q} \leq C \leq \overline{Q}\big\}$$

is a closed, bounded and equicontinuous subset of the topological space of all continuous functions on $[0, 1]^d$, equipped with the uniform metric. Hence, it follows from the Arzelà-Ascoli Theorem that $\mathcal{C}_b$ is compact. Moreover, $\mathcal{C}_b$ and $\mathcal{R}^2$ are convex sets. On the other hand, for all $h, g \in \mathcal{R}$ the map $f\big(\cdot\,, (h, g)\big)$ is continuous w.r.t. the uniform convergence of copulas since we assume φ to be bounded and continuous. Furthermore, we have that $f\big(\cdot\,, (h, g)\big)$ is convex on $\mathcal{C}^d$. Also, for all $C \in \mathcal{C}^d$ it holds that $f(C, \cdot)$ is linear on $\mathcal{R}^2$. To verify (3.11), assume that $\underline{Q}(h) \leq C(h) \leq \overline{Q}(h)$ does not hold for one $h \in \mathcal{R}$, i.e. let w.l.o.g. $C(h) < \underline{Q}(h)$, then for each $\alpha > 0$ it follows that

$$\big(\underline{Q}(\alpha h) - C(\alpha h)\big) = \alpha\big(\underline{Q}(h) - C(h)\big) > 0$$

and thus, by scaling α, the supremum is ∞ and C can be disregarded in the infimum in (3.10). Hence, it holds that

$$\underline{Q}(h) \leq C(h) \leq \overline{Q}(h), \text{for all } h \in \mathcal{R}.$$

This entails that $\underline{Q}(h) - C(h) \leq 0$ and $-(\overline{Q}(g) - C(g)) \leq 0$ for all $(g, h) \in \mathcal{R}^2$ and thus the supremum is attained for $h, g \equiv 0$. Finally, (3.12) holds due to the fact that $\underline{Q}(h) \leq C(h) \leq \overline{Q}(h)$ for all $h \in \mathcal{R}$ implies

$$\underline{Q}(F_1(x_1), ..., F_d(x_d)) \leq C(F_1(x_1), ..., F_d(x_d)) \leq \overline{Q}(F_1(x_1), ..., F_d(x_d))$$

for all $(x_1, ..., x_d) \in \mathbb{R}^d$ and $\mathbb{E}_C[\varphi] = \mathbb{E}_{C'}[\varphi]$ for all copulas C and C' with $C(F_1(x_1), ..., F_d(x_d)) = C'(F_1(x_1), ..., F_d(x_d))$.

We proceed by relaxing the condition of φ being bounded and continuous. So let φ merely be lower semicontinuous. We can w.l.o.g. assume that $\varphi \geq 0$ as otherwise there exist, due to condition (3.5), functions $g_1, ...g_d$ with $\varphi + \sum_{i=1}^d g_i \geq 0$ and

$$\overline{P}_\varphi = \overline{P}_{\varphi+\sum_{i=1}^d g_i} - \sum_{i=1}^d \mathbb{E}_i[g_i].$$

Now, since φ is lower semicontinuous there exists a sequence of positive, bounded, continuous functions $\varphi_1 \leq \varphi_2 \leq \cdots$ with $\varphi = \lim_n \varphi_n$ pointwise and $\underline{P}_{\varphi_n} \leq \underline{P}_\varphi$. Furthermore, due to the compactness of $\mathcal{C}_b$ there exist optimizers $C_1, C_2, ...$ of $\underline{P}_{\varphi_1}, \underline{P}_{\varphi_2}, ...$ and we can, by passing to a subsequence, assume that $C_1, C_2, ...$ converges to some $C^* \in \mathcal{C}_b$. Then it follows by monotone convergence that

$$\underline{P}_\varphi \leq \mathbb{E}_{C^*}[\varphi] = \lim_n \mathbb{E}_{C^*}[\varphi_n] = \lim_n \lim_j \mathbb{E}_{C_j}[\varphi_n] \leq \lim_j \mathbb{E}_{C_j}[\varphi_j] = \lim_j \underline{P}_{\varphi_j} = \lim_j \underline{D}_{\varphi_j} = \underline{D}_\varphi,$$

where the last equality is due to $\underline{D}_{\varphi_j} \leq \underline{D}_\varphi \leq \underline{P}_\varphi$ for $j \in \mathbb{N}$.

Lastly, we note that the optimizers for the generalized Fréchet functionals are attained due to the compactness of $\mathcal{C}_b$ which completes the proof. □

Remark 3.3. Assuming the existence of a copula $C \in \mathcal{C}^d$ with $\underline{Q} \leq C \leq \overline{Q}$ in Theorem 3.2 rules out the degenerate situation where no probabilistic model exists which is compatible with the prescribed information. Verifying this assumption however is a delicate task in general. The existence of a copula C with $\underline{Q} \leq C$ follows immediately from the fact that the upper Fréchet–Hoeffding bound M_d is a copula and hence $\underline{Q} \leq M_d$. The difficulty thus lies in verifying $C \leq \overline{Q}$ which fails e.g. when $\overline{Q} = W_d$ and $d > 2$, where W_d is the lower Fréchet–Hoeffding bound given in (2.4). Nevertheless, when $\underline{Q}$ and $\overline{Q}$ are improved Fréchet–Hoeffding bounds it is often straight-forward to verify that $\{C \in \mathcal{C}^d : \underline{Q} \leq C \leq \overline{Q}\}$ is not empty. ♦

Remark 3.4. The duality result with additional dependence information in Theorem 3.2 was also developed, in parallel and using a completely different proof, in Bartl, Kupper, Lux, and Papapantoleon (2017, Theorem 2.2). ♦

The following counter-example shows that the dual optimizers are not attained in general.

Example 3.5. Consider the case $d = 2$ and let F_1 and F_2 be uniform marginal distribution on $[0, 1]$. Moreover, let $\underline{Q}(u_1, u_2) = \overline{Q}(u_1, u_2) = \Pi(u_1, u_2) = u_1 u_2$ for all $(u_1, u_2) \in [0, 1]^2$ and consider $\varphi\colon \mathbb{R}^2 \to \mathbb{R}\colon (u_1, u_2) \mapsto \mathbb{1}_{\psi(u_1,u_2)<1}$ where $\psi(u_1, u_2) = \sqrt{u_1^2 + u_2^2}$, i.e. φ is the indicator function of the circular segment of the unit circle on $[0, 1]^2$. It then follows from $\underline{Q} = \overline{Q} = \Pi$, that

$$\underline{P}_\varphi = \overline{P}_\varphi = \int_{[0,1]^2} \mathbb{1}_{\sqrt{u_1^2+u_2^2}<1} \, \mathrm{d}u_1 \mathrm{d}u_2 = \frac{\pi}{4}.$$

Now, assume the dual optimizer for $\underline{D}_\varphi$ is attained. Then it is of the form

$$f^* := h - g + f_1 + f_2 = \sum_{n=1}^{k} \alpha_n \Lambda_{\mathbf{u}^n} + \sum_{n=1}^{m} -\beta_n \Lambda_{\mathbf{v}^n} + f_1 + f_2$$

and since $f^*(u_1, u_2) \leq \mathbb{1}_{\psi(u_1,u_2)<1}$ for all $(u_1, u_2) \in [0,1]^2$ and $\mathbb{E}_\Pi[f^*] = \frac{\pi}{4}$, we have that

$$f^*(u_1, u_2) = (h - g + f_1 + f_2)(u_1, u_2) = \mathbb{1}_{\psi(u_1,u_2)<1} \quad \lambda\text{-a.s.} \tag{3.13}$$

Moreover, we can assume w.l.o.g. that $f_1 \equiv f_2 \equiv 0$ λ-a.s. since it follows from equation (3.13) that

$$(h - g)(u_1, 1) = \mathbb{1}_{\psi(u_1,1)<1} - f_1(u_1) - f_2(1) = -f_1(u_1) - c \quad \lambda\text{-a.s.}$$

where the last equality is due to $\mathbb{1}_{\psi(u_1,1)<1} = 0$ λ-a.s. and $f_2(1) =: c$. Now, by the same argument it follows that

$$(h - g)(1, u_2) = -f_2(u_2) - c' \quad \lambda\text{-a.s.},$$

and thus we obtain

$$(h - g)(u_1, u_d) = \left(\sum_{n=1}^{k} \alpha_n \Lambda_{\mathbf{u}^n} + \sum_{n=1}^{m} -\beta_n \Lambda_{\mathbf{v}^n} \right)(u_1, u_2) = \mathbb{1}_{\psi(u_1,u_2)<1} \quad \lambda\text{-a.s.}$$

This however, corresponds to a construction of the indicator function of the circular segment by a finite number of rectangular indicator functions which contradicts the impossibility of the squaring of the circle. $\Diamond$

4. A reduction scheme to compute bounds on the Value-at-Risk

The dual characterizations of $\underline{P}_\varphi$ and $\overline{P}_\varphi$ in Section 3 lend themselves to the development of a scheme to compute VaR estimates, accounting for an upper and a lower bound on the copula of the risks. In general, the dual problems do not admit closed form solutions except when $\underline{Q} = W_d$, $\overline{Q} = M_d$ with homogeneous marginals $F_1 = \cdots = F_d$ fulfilling additional constraints; c.f. Wang and Wang (2011) and Puccetti and Rüschendorf (2013). We therefore develop in this section a scheme that corresponds to an optimization over a tractable subset of admissible functions for the duals $\underline{D}_\varphi$ and $\overline{D}_\varphi$ that produces narrow VaR bounds. Furthermore, we show that the scheme produces asymptotically sharp bounds in the certainty limit, i.e. when $\underline{Q}$ and $\overline{Q}$ converge to some copula C.

4.1. A reduction scheme for $\underline{D}_\varphi$

Consider the function $\varphi(x_1, ..., x_d) = \mathbb{1}_{\psi(x_1,...,x_d)<s}$ for componentwise increasing $\psi\colon \mathbb{R}^d \to \mathbb{R}$ and recall from (2.2) that our generalized Fréchet functional of interest reads

$$\underline{P}_\varphi := \inf \left\{ \mathbb{E}_C[\varphi] \colon C \in \mathcal{C}^d, \underline{Q} \leq C \leq \overline{Q} \right\},$$

for quasi-copulas $\underline{Q}$ and $\overline{Q}$, whereas the corresponding dual problem is given in (3.2) by

$$\underline{D}_\varphi = \sup\Bigg\{\underline{Q}(h) - \overline{Q}(g^-) + \sum_{i=1}^d \mathbb{E}_i[f_i]\colon f_i \in \mathcal{L}(F_i),\ i = 1, ..., d,$$
$$h, g \in \mathcal{R} \text{ s.t. } h - g^- + \sum_{i=1}^d f_i \le \varphi\Bigg\}.$$

In the following, we identify admissible functions for the dual $\underline{D}_\varphi$ by $(d+2)$-tuples in the class

$$\underline{\mathcal{A}} := \left\{(h, g, f_1, ..., f_d)\colon f_i \in \mathcal{L}(F_i),\ i = 1, ..., d,\ h, g \in \mathcal{R} \text{ s.t. } h - g^- + \sum_{i=1}^d f_i \le \varphi\right\}$$

and for each admissible tuple the corresponding value of the objective function amounts to

$$\underline{Q}(h) - \overline{Q}(g^-) + \sum_{i=1}^d \mathbb{E}_i[f_i].$$

Regarding the improved standard bounds in Embrechts et al. (2003) and Embrechts and Puccetti (2006), we note that when the copula C of $\mathbf{X}$ is bounded from below by $\underline{Q}$, i.e. $\underline{Q} \le C$, then the lower improved standard bound is given by

$$\mathbb{E}_C[\mathbb{1}_{\psi(\mathbf{X})<s}] \ge \sup_{u_1,\dots,u_{d-1}\in\mathbb{R}} \underline{Q}\big(F_1(u_1), \dots, F_{d-1}(u_{d-1}), F_d^-(\psi^*_{u_{-d}}(s))\big) = \underline{m}_{\underline{Q},\psi}(s),$$

where $\psi^*_{u_{-d}}(s) = \sup\{u \in \mathbb{R}\colon u_1 + \cdots + u_{d-1} + u < s\}$. The bound on the expectation corresponds, in the case of continuous marginals, to the maximization of $\underline{Q}(h)$ over functions $h = \Lambda_{\mathbf{u}} \in \mathcal{R}$ with $\mathbf{u} \in \big\{\big(u_1, ..., u_{d-1}, \psi^*_{u_{-d}}(s)\big)\colon (u_1, ...u_{d-1}) \in \mathbb{R}^{d-1}\big\}$. Hence, $\underline{m}_{\underline{Q},\psi}(s)$ can be viewed as an optimization over a – rather small – subset of admissible elements in $\underline{\mathcal{A}}$, i.e. tuples of the form $(h, 0, ..., 0) \in \underline{\mathcal{A}}$.

Leveraging this observation, we develop an optimization scheme over a larger subset of admissible functions. To this end, let us first consider admissible $(h, g, f_1, ..., f_d)$ with

- **(A1)** $f_1, ..., f_d \equiv 0$,
- **(A2)** $h, g \in \mathcal{R}^r$, where

$$\mathcal{R}^r := \left\{\sum_{n=1}^k \Lambda_{\mathbf{u}^n} : k \in \mathbb{N}, \mathbf{u}_1, ..., \mathbf{u}_k \in \mathcal{U}_\psi(s)\right\},$$

and $\mathcal{U}_\psi(s) = \{(x_1, ..., x_d) \in \mathbb{R}^d\colon \psi(x_1, ..., x_d) < s\}$. We thus obtain a set of admissible functions given by

$$\underline{\mathcal{A}}^r := \big\{(h, g)\colon h, g \in \mathcal{R}^r \text{ s.t. } h - g^- \le \varphi\big\}.$$

Note, that – by abuse of terminology – $\underline{\mathcal{A}}^r$ can be viewed as a subset of $\underline{\mathcal{A}}$, where elements $(h, g) \in \underline{\mathcal{A}}^r$ are identified with $(h, g, 0, ..., 0) \in \underline{\mathcal{A}}$ for $h, g \in \mathcal{R}^r$. The optimization over the subset $\underline{\mathcal{A}}^r$ remains however intractable due to the constraint $h - g^- \leq \varphi$. Moreover, since $\mathcal{R}^r$ consists of sums of the form $\sum_{n=1}^k \Lambda_{\mathbf{u}^n}$ for $k \in \mathbb{N}$, optimizing over $\underline{\mathcal{A}}^r$ requires a truncation of the variable k. An appropriate choice for such a truncation is not obvious. We therefore proceed with the development of an unconstrained optimization scheme over a finite number of elements in $\mathcal{U}_\psi(s)$. An informal description and illustration of the scheme and the idea of the proof is provided in Section 5. For notational ease, we introduce the notion of multisets (c.f. Definition 2 in Syropoulos (2001)).

Definition 4.1. Let $\mathcal{B}$ be some set. A multiset over $\mathcal{B}$ is a pair $\langle \mathcal{B}, f \rangle$ where $f \colon \mathcal{B} \to \mathbb{N}$ and f is called *multiplicity function*.

Remark 4.2. Multisets generalize the notion of a set so as to allow for finite but multiple occurrences of elements. By the conventional notion of a set we have that $\mathcal{B} := \{1, 1, 2\} = \{1, 2\}$. Using the notion multisets we refer to $\{1, 1, 2\}$ as $\langle \mathcal{B}, f \rangle$ with $f(1) = 2$ and $f(2) = 1$. The multiplicity function f hence counts the number of occurences of each element of $\mathcal{B}$. ♦

Our scheme is based on the following inclusion-exclusion principle for multisets.

Lemma 4.3 (Multiset inclusion-exclusion principle). *Let $B_1, ..., B_k \subset \mathbb{R}^d$ and define for $m = 1, ..., k$ the multisets*

$$\begin{aligned} &\langle \mathcal{B}^o, f^o \rangle, \quad \mathcal{B}^o := \{B_{i_1} \cap \cdots \cap B_{i_m} \colon 1 \leq i_1 < \cdots < i_m \leq k,\ m \text{ odd}\} \\ &\langle \mathcal{B}^e, f^e \rangle, \quad \mathcal{B}^e := \{B_{i_1} \cap \cdots \cap B_{i_m} \colon 1 \leq i_1 < \cdots < i_m \leq k,\ m \text{ even}\} \end{aligned} \tag{4.1}$$

where

$$f^o(B) = |\{(i_1, ..., i_m) \colon 0 \leq i_1 < \cdots < i_m \leq k,\ m \text{ odd},\ B = B_{i_1} \cap \cdots \cap B_{i_m}\}|,$$

for $B \in \mathcal{B}^o$ and f^e is defined analogously. Then

$$\mathbb{1}_{B_1 \cup \cdots \cup B_k} = \sum_{B \in \mathcal{B}^o} \big(f^o(B) - f^e(B)\big)^+ \mathbb{1}_B - \sum_{B \in \mathcal{B}^e} \big(f^e(B) - f^o(B)\big)^+ \mathbb{1}_B.$$

Proof. Applying the classical inclusion-exclusion principle (see e.g. Loera, Hemmecke, and Köppe (2013, Lemma 6.1.2)) to $\mathbb{1}_{B_1 \cup \cdots \cup B_k}$ yields

$$\mathbb{1}_{B_1 \cup \cdots \cup B_k} = \sum_{B \in \mathcal{B}^o} f^o(B)\, \mathbb{1}_B - \sum_{B \in \mathcal{B}^e} f^e(B)\, \mathbb{1}_B.$$

Then, by rearranging the terms and using the fact that $f^e(B) = 0$ when $B \in \mathcal{B}^o \setminus \mathcal{B}^e$ and $f^o(B) = 0$ when $B \in \mathcal{B}^e \setminus \mathcal{B}^o$ we obtain

$$\begin{aligned} \sum_{B \in \mathcal{B}^o} f^o(B)\, \mathbb{1}_B - \sum_{B \in \mathcal{B}^e} f^e(B) \mathbb{1}_B &= \sum_{B \in \mathcal{B}^o \setminus \mathcal{B}^e} f^o(B)\, \mathbb{1}_B - \sum_{B \in \mathcal{B}^e \setminus \mathcal{B}^o} f^e(B)\, \mathbb{1}_B \\ &\quad + \sum_{B \in \mathcal{B}^o \cap \mathcal{B}^e} \big(f^o(B) - f^e(B)\big)\, \mathbb{1}_B \end{aligned}$$

The statement then follows from

$$\big(f^o(B) - f^e(B)\big) = \big(f^o(B) - f^e(B)\big)^+ - \big(f^e(B) - f^o(B)\big)^+.$$

which completes the proof. □

Remark 4.4. Lemma 4.3 establishes a non-redundant version of the classical inclusion-exclusion principle. To illustrate this, consider $B_1, B_2, B_3 \in \mathbb{R}^d$ such that $B_1 \cap B_2 \cap B_3 = B_1 \cap B_2$ and $B_1 \neq B_2$. Then applying the classical inclusion-exclusion principle to $B_1 \cup B_2 \cup B_3$ yields

$$\mathbb{1}_{B_1\cup B_2\cup B_3} = \mathbb{1}_{B_1} + \mathbb{1}_{B_2} + \mathbb{1}_{B_3} - \mathbb{1}_{B_1\cap B_2} - \mathbb{1}_{B_1\cap B_3} - \mathbb{1}_{B_2\cap B_3} + \mathbb{1}_{B_1\cap B_2\cap B_3},$$

where the terms $-\mathbb{1}_{B_1\cap B_2}$ and $+\mathbb{1}_{B_1\cap B_2\cap B_3}$ cancel each other out. This superfluous subtraction and addition of terms is avoided using the multisets $\langle \mathcal{B}^o, f^o\rangle$ and $\langle \mathcal{B}^e, f^e\rangle$ as in Lemma 4.3. Due to $f^o(B_1 \cap B_2) = f^e(B_1 \cap B_2 \cap B_3)$ we then obtain

$$\mathbb{1}_{B_1\cup B_2\cup B_3} = \mathbb{1}_{B_1} + \mathbb{1}_{B_2} + \mathbb{1}_{B_3} - \mathbb{1}_{B_1\cap B_3} - \mathbb{1}_{B_2\cap B_3},$$

and thus a more parsimonious representation of $\mathbb{1}_{B_1\cup B_2\cup B_3}$. This is a relevant improvement of the usual inclusion-exclusion principle when two-sided bounds on the distribution are available. ♦

In the following we denote the componentwise minimum of vectors $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathbb{R}^d$ by

$$\min(\mathbf{u}^1, ..., \mathbf{u}^k) = \Big(\min_{n=1,...,k}\{u_1^n\}, ..., \min_{n=1,...,k}\{u_d^n\}\Big).$$

Moreover, we define the sets

$$\begin{aligned}\mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k) &:= \{\min(\mathbf{u}^{i_1}, ..., \mathbf{u}^{i_m})\colon 1 \le i_1 < \cdots < i_m \le k \;\; m \text{ odd}\}\\ \mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^k) &:= \{\min(\mathbf{u}^{i_1}, ..., \mathbf{u}^{i_m})\colon 1 \le i_1 < \cdots < i_m \le k \;\; m \text{ even}\}.\end{aligned} \tag{4.2}$$

We refer to the multiplicity function

$$l^o(\mathbf{u}) := |\{(i_1, ..., i_m)\colon 0 \le i_1 < \cdots < i_m \le k,\; m \text{ odd},\; \mathbf{u} = \min(\mathbf{u}^{i_1}, ..., \mathbf{u}^{i_m})\}|$$

for $\mathbf{u} \in \mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k)$ as the *multiplicity function associated to* $\mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k)$ and define the multiplicity function associated to $\mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^k)$, denoted by l^e, analogously. The normalization of a vector $\mathbf{u} \in \mathbb{R}^d$ by the marginals $F_1, ..., F_d$ is denoted by $F(\mathbf{u}) := \big(F_1(u_1), ..., F_d(u_d)\big)$ as well as the left-continuous version $F^-(\mathbf{u}) := \big(F_1^-(u_1), ..., F_d^-(u_d)\big)$. Finally, for $\boldsymbol{\varepsilon} := (\varepsilon, ..., \varepsilon) \in \mathbb{R}^d$ and $\mathbf{u} \in \mathbb{R}^d$ we denote $\mathbf{u} + \boldsymbol{\varepsilon} = (u_1 + \varepsilon, ..., u_d + \varepsilon)$.

Lemma 4.5. *Let $k \in \mathbb{N}$ and $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_\psi(s)$. Define the functions*

$$h := \sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \Lambda_{\mathbf{u}}; \quad g_{\boldsymbol{\varepsilon}} := \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \Lambda_{\mathbf{u}+\boldsymbol{\varepsilon}},$$

for l^o and l^e being the multiplicity functions associated to $\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^k)$ and $\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)$ respectively. Then for every $\boldsymbol{\varepsilon} > \mathbf{0}$ it holds that $(h, g_{\boldsymbol{\varepsilon}})$ is admissible for the dual problem $\underline{D}_\varphi$, i.e. $(h, g_{\boldsymbol{\varepsilon}}) \in \underline{\mathcal{A}}^r$, and the value of the objective function is given by

$$\sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \underline{Q}\big(F(\mathbf{u})\big) - \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \overline{Q}\big(F^-(\mathbf{u}+\boldsymbol{\varepsilon})\big).$$

Proof. It suffices to show that $h - g_{\boldsymbol{\varepsilon}}^- \leq \varphi$ for any $\boldsymbol{\varepsilon} > \mathbf{0}$. Recalling the notion of the sublevel set

$$\mathcal{U}_\psi(s) = \big\{(u_1, ...u_d) \in \mathbb{R}^d \colon \psi(u_1, ..., u_d) < s\big\},$$

we have for every $(u_1, ..., u_d) \in \mathcal{U}_\psi(s)$ that

$$\big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon x_1 \leq u_1, ..., x_d \leq u_d\big\} \subset \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon \psi(x_1, \cdots, x_d) < s\big\},$$

due to the fact that ψ is increasing in each coordinate.

Hence, for $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_\psi(s)$ and $B_n := \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon x_1 \leq u_1^n, ..., x_d \leq u_d^n\big\}$ for $n = 1, ..., k$ we have that

$$\bigcup_{n=1}^k B_n \subset \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon \psi(x_1, \cdots, x_d) < s\big\}.$$

Now applying the inclusion-exclusion principle for multisets (Lemma 4.3) to $\bigcup_{n=1}^k B_n$ we obtain

$$\mathbb{1}_{B_1\cup\cdots\cup B_k} = \sum_{B\in\mathcal{B}^o} \big(f^o(B) - f^e(B)\big)^+ \mathbb{1}_B - \sum_{B\in\mathcal{B}^e} \big(f^e(B) - f^o(B)\big)^+ \mathbb{1}_B,$$

where $\mathcal{B}^o$ and $\mathcal{B}^e$ are as in (4.1) and f^o, f^e are the respective multiplicity functions. Moreover, we have for $\bigcap_{l=1}^m B_{n_l} \in \mathcal{B}^o \cup \mathcal{B}^e$ that

$$\begin{aligned}\bigcap_{l=1}^m B_{n_l} &= \bigcap_{l=1}^m \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon x_1 \leq u_1^{n_l}, ..., x_d \leq u_d^{n_l}\big\} \\ &= \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon x_1 \leq \min(u_1^{n_1}, ..., u_1^{n_m}), ..., x_d \leq \min(u_d^{n_1}, ..., u_d^{n_m})\big\} \\ &= \{\mathbf{x} \in \mathbb{R}^d \colon \mathbf{x} \leq \min(\mathbf{u}^{n_1}, ..., \mathbf{u}^{n_m})\},\end{aligned}$$

and thus $\mathbb{1}_{B_{n_1}\cap\cdots\cap B_{n_m}} = \Lambda_{\min(\mathbf{u}^{n_1},...,\mathbf{u}^{n_m})}$ for $\min(\mathbf{u}^{n_1}, ..., \mathbf{u}^{n_m}) \in \mathcal{M}^o \cup \mathcal{M}^e$. Also, if $B = \bigcap_{l=1}^m B_{n_l} \in \mathcal{B}^o$ we have that

$$f^o(B) = l^o(\min(\mathbf{u}^{n_1}, ..., \mathbf{u}^{n_m}))$$

and $f^e(B) = l^e(\min(\mathbf{u}^{n_1}, ..., \mathbf{u}^{n_m}))$ for $B = \bigcap_{l=1}^m B_{n_l} \in \mathcal{B}^e$. In particular, it follows for any

$\boldsymbol{\varepsilon} > \mathbf{0}$ that

$$
\begin{aligned}
h(\mathbf{x}) - g_{\boldsymbol{\varepsilon}}^-(\mathbf{x}) &\leq h(\mathbf{x}) - g_{\mathbf{0}}(\mathbf{x}) \\
&\leq \sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,\dots,\mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \Lambda_{\mathbf{u}}(\mathbf{x}) - \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,\dots,\mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \Lambda_{\mathbf{u}}(\mathbf{x}) \\
&= \sum_{B\in\mathcal{B}^o} \big(f^o(B) - f^e(B)\big)^+ \mathbb{1}_B(\mathbf{x}) - \sum_{B\in\mathcal{B}^e} \big(f^e(B) - f^o(B)\big)^+ \mathbb{1}_B(\mathbf{x}) \\
&= \mathbb{1}_{B_1\cup\cdots\cup B_k}(\mathbf{x}) \leq \mathbb{1}_{\psi(\mathbf{x})<s} = \varphi(\mathbf{x})
\end{aligned}
$$

for all $\mathbf{x} \in \mathbb{R}^d$ and so $(h, g_{\boldsymbol{\varepsilon}}) \in \underline{\mathcal{A}}^r$ which completes the proof. □

We are now in the position to state the reduced optimization problem for $\underline{D}_\varphi$.

Theorem 4.6 (Reduced lower dual bound with two-sided dependence information). *Let $\varphi(x_1, \dots, x_d) = \mathbb{1}_{\psi(x_1,\dots,x_d)<s}$ where $\psi\colon \mathbb{R}^d \to \mathbb{R}$ is increasing in each coordinate and let*

$$
\begin{aligned}
\underline{D}_\varphi(k) := \sup\Bigg\{ &\sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,\dots,\mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \underline{Q}\big(F(\mathbf{u})\big) \\
&- \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,\dots,\mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \overline{Q}\big(F(\mathbf{u})\big) \colon \mathbf{u}^1, \dots, \mathbf{u}^k \in \mathcal{U}_\psi(s) \Bigg\},
\end{aligned} \tag{4.3}
$$

where l^o and l^e are the canonical multiplicity functions associated to $\mathcal{M}^o(\mathbf{u}^1, \dots, \mathbf{u}^k)$ and $\mathcal{M}^e(\mathbf{u}^1, \dots, \mathbf{u}^k)$ respectively. Then

$$\underline{D}_\varphi(k) \leq \underline{D}_\varphi(k+1) \leq \dots \leq \underline{D}_\varphi.$$

Proof. We first show that $\underline{D}_\varphi(k) \leq \underline{D}_\varphi(k+1)$ for $k \in \mathbb{N}$. Therefore note that when $\mathbf{u}^k = \mathbf{u}^{k+1} \in \mathcal{U}_\psi(s)$ it follows that

$$\mathcal{M}^o(\mathbf{u}^1, \dots, \mathbf{u}^{k+1}) \cup \mathcal{M}^e(\mathbf{u}^1, \dots, \mathbf{u}^{k+1}) = \mathcal{M}^o(\mathbf{u}^1, \dots, \mathbf{u}^k) \cup \mathcal{M}^e(\mathbf{u}^1, \dots, \mathbf{u}^k)$$

and defining

$$
\begin{aligned}
\mathcal{M}^o|_{\,i_{m-1}\neq k, i_m = k+1} &:= \{\min(\mathbf{u}^{i_1}, \dots, \mathbf{u}^{i_m}) \colon 1 \leq i_1 < \cdots < i_m \leq k+1,\ i_{m-1} \neq k, \\
&\qquad\qquad i_m = k+1,\ m \text{ odd}\} \\
\mathcal{M}^o|_{\,i_{m-1}= k, i_m = k+1} &:= \{\min(\mathbf{u}^{i_1}, \dots, \mathbf{u}^{i_m}) \colon 1 \leq i_1 < \cdots < i_m \leq k+1\ i_{m-1} = k, \\
&\qquad\qquad i_m = k+1,\ m \text{ odd}\},
\end{aligned}
$$

it holds

$$\mathcal{M}^o(\mathbf{u}^1, \dots, \mathbf{u}^{k+1}) = \mathcal{M}^o(\mathbf{u}^1, \dots, \mathbf{u}^k) \cup \mathcal{M}^o|_{\,i_{m-1}\neq k, i_m = k+1} \cup \mathcal{M}^o|_{\,i_{m-1}= k, i_m = k+1}.$$

Moreover, noting that $\mathcal{M}^o|_{\,i_{m-1}\neq k, i_m=k+1} = \mathcal{M}^o|_{\,i_{m-1}=k, i_m=k+1}$ and

$$\begin{aligned}\mathcal{M}^o|_{\,i_{m-1}\neq k, i_m=k+1} &= \{\min(\mathbf{u}^{i_1},\dots,\mathbf{u}^{i_m})\colon 1\le i_1<\cdots<i_m\le k+1\ \ i_{m-1}=k,\\ &\qquad\qquad\qquad\qquad\qquad\qquad i_m=k+1,\ m\text{ even}\}\\ \mathcal{M}^o|_{\,i_{m-1}=k, i_m=k+1} &= \{\min(\mathbf{u}^{i_1},\dots,\mathbf{u}^{i_m})\colon 1\le i_1<\cdots<i_m\le k+1\ \ i_{m-1}\neq k,\\ &\qquad\qquad\qquad\qquad\qquad\qquad i_m=k+1,\ m\text{ even}\},\end{aligned}$$

it follows, by straight forward computation, that

$$(l^o_{k+1}(\mathbf{u})-l^e_{k+1}(\mathbf{u}))^+ = (l^o_k(\mathbf{u})-l^e_k(\mathbf{u}))^+,\quad \text{for all } \mathbf{u}\in\mathcal{M}^o|_{\,i_{m-1}\neq k, i_m=k+1},$$

where l^o_j, l^e_j are the canonical multiplicity functions associated to $\mathcal{M}^o(\mathbf{u}^1,\dots,\mathbf{u}^j)$ and $\mathcal{M}^e(\mathbf{u}^1,\dots,\mathbf{u}^j)$ respectively for $j=k,k+1$. Applying a similar argument to the set $\mathcal{M}^e(\mathbf{u}^1,\dots,\mathbf{u}^{k+1})$ yields that

$$\begin{aligned}&\sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,\dots,\mathbf{u}^{k+1})}\big(l^o_{k+1}(\mathbf{u})-l^e_{k+1}(\mathbf{u})\big)^+\,\underline{Q}\big(F(\mathbf{u})\big)\\ &\qquad\qquad-\sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,\dots,\mathbf{u}^{k+1})}\big(l^e_{k+1}(\mathbf{u})-l^o_{k+1}(\mathbf{u})\big)^+\,\overline{Q}\big(F(\mathbf{u})\big)\\ &=\sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,\dots,\mathbf{u}^{k})}\big(l^o_{k}(\mathbf{u})-l^e_{k}(\mathbf{u})\big)^+\,\underline{Q}\big(F(\mathbf{u})\big)\ -\sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,\dots,\mathbf{u}^{k})}\big(l^e_{k}(\mathbf{u})-l^o_{k}(\mathbf{u})\big)^+\,\overline{Q}\big(F(\mathbf{u})\big).\end{aligned}$$

This implies that the values of the objective functions coincide when $\mathbf{u}^k=\mathbf{u}^{k+1}$ and hence it follows in particular that $\underline{D}_\varphi(k)\le\underline{D}_\varphi(k+1)$ for all $k\in\mathbb{N}$.

Furthermore, the inequality $\underline{D}_\varphi(k)\le\underline{D}_\varphi$ for all $k\in\mathbb{N}$ follows by an application of Lemma 4.5 to $(h,g_{\boldsymbol{\varepsilon}})$ where $\boldsymbol{\varepsilon}>\mathbf{0}$ and

$$h=\sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,\dots,\mathbf{u}^k)}\big(l^o(\mathbf{u})-l^e(\mathbf{u})\big)^+\,\Lambda_{\mathbf{u}};\quad g_{\boldsymbol{\varepsilon}}=\sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,\dots,\mathbf{u}^k)}\big(l^e(\mathbf{u})-l^o(\mathbf{u})\big)^+\,\Lambda_{\mathbf{u}+\boldsymbol{\varepsilon}}.$$

This yields that $(h,g_{\boldsymbol{\varepsilon}})\in\underline{\mathcal{A}}^r$ and the respective value of the objective function is given by

$$\sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,\dots,\mathbf{u}^k)}\big(l^o(\mathbf{u})-l^e(\mathbf{u})\big)^+\,\underline{Q}\big(F(\mathbf{u})\big)-\sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,\dots,\mathbf{u}^k)}\big(l^e(\mathbf{u})-l^o(\mathbf{u})\big)^+\,\overline{Q}\big(F^-(\mathbf{u}+\boldsymbol{\varepsilon})\big).$$

In particular, we have for all $k\in\mathbb{N}$ that

$$\begin{aligned}&\sup\Bigg\{\sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,\dots,\mathbf{u}^k)}\big(l^o(\mathbf{u})-l^e(\mathbf{u})\big)^+\,\underline{Q}\big(F(\mathbf{u})\big)\\ &-\sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,\dots,\mathbf{u}^k)}\big(l^e(\mathbf{u})-l^o(\mathbf{u})\big)^+\,\overline{Q}\big(F^-(\mathbf{u}+\boldsymbol{\varepsilon})\big)\colon\boldsymbol{\varepsilon}>\mathbf{0};\ \mathbf{u}^1,\dots,\mathbf{u}^k\in\mathcal{U}_\psi(s)\Bigg\}=:c^*\le\underline{D}_\varphi.\end{aligned}\tag{4.4}$$

Moreover, it holds for all $\boldsymbol{\varepsilon} > \mathbf{0}$ that

$$\sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,\dots,\mathbf{u}^k)} \left(l^o(\mathbf{u}) - l^e(\mathbf{u})\right)^+ \underline{Q}\big(F(\mathbf{u})\big) - \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,\dots,\mathbf{u}^k)} \left(l^e(\mathbf{u}) - l^o(\mathbf{u})\right)^+ \overline{Q}\big(F^-(\mathbf{u}+\boldsymbol{\varepsilon})\big)$$
$$\leq \sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,\dots,\mathbf{u}^k)} \left(l^o(\mathbf{u}) - l^e(\mathbf{u})\right)^+ \underline{Q}\big(F(\mathbf{u})\big) - \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,\dots,\mathbf{u}^k)} \left(l^e(\mathbf{u}) - l^o(\mathbf{u})\right)^+ \overline{Q}\big(F(\mathbf{u})\big)$$

as well as $\lim_{\boldsymbol{\varepsilon}\to\mathbf{0}} \overline{Q}\big(F^-(\mathbf{u}+\boldsymbol{\varepsilon})\big) = \overline{Q}\big(F(\mathbf{u})\big)$ due to the Lipschitz continuity of $\overline{Q}$. Hence using (4.4) it follows that

$$c^* = \sup\Bigg\{ \sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,\dots,\mathbf{u}^k)} \left(l^o(\mathbf{u}) - l^e(\mathbf{u})\right)^+ \underline{Q}\big(F(\mathbf{u})\big)$$
$$- \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,\dots,\mathbf{u}^k)} \left(l^e(\mathbf{u}) - l^o(\mathbf{u})\right)^+ \overline{Q}\big(F(\mathbf{u})\big) \colon \mathbf{u}^1,\dots,\mathbf{u}^k \in \mathcal{U}_\psi(s) \Bigg\}$$

which completes the proof. □

Theorem 4.6 establishes a tractable optimization problem to compute a lower bound on $\underline{D}_\varphi$ and thus also on $\underline{P}_\varphi$. The optimization takes place over vectors in the sublevel set $\mathcal{U}_\psi(s)$ and the trade-off between the computational effort and the quality of the bound is moderated by the variable k. For fixed k, $\underline{D}_\varphi(k)$ amounts to a $k \cdot d$ dimensional optimization that can be solved with standard optimization packages. Note, that most mathematical programming environments also provide efficient built-in procedures to compute the multiplicity functions l^o and l^k[2].

Remark 4.7. From Theorem 4.6 it is evident that the complexity of the maximization in (4.3) can be further reduced by restricting the optimization in a suitable way to a smaller subset of vectors in $\mathcal{U}_\psi(s)$. This observation is key to the development of heuristics to compute bounds in large dimensions. The development of heuristics which reduce complexity while still providing reasonably narrow bounds on the expectation in high dimensions is subject to our ongoing research. ♦

4.2. A reduction scheme for $\overline{D}_\varphi$

We proceed with the development of a similar reduction scheme based on the dual $\overline{D}_\varphi$. Recall from (3.3) that

$$\overline{D}_\varphi = \inf\Bigg\{ \overline{Q}(h^-) - \underline{Q}(g) + \sum_{i=1}^d \mathbb{E}_i[f_i] \colon f_i \in \mathcal{L}(F_i),\ i = 1,\dots,d,$$
$$h, g \in \mathcal{R} \text{ s.t. } h^- - g + \sum_{i=1}^d f_i \geq \varphi \Bigg\}.$$

[2] The programming language MATLAB provides e.g. the functions *nchoosek* to generate the index set used in the definition of $\mathcal{M}^o$ and $\mathcal{M}^e$ and duplicate vectors can be removed using the matrix command *unique*

We refer to the class of admissible functions for $\overline{D}_\varphi$ by

$$\overline{\mathcal{A}} := \left\{(h, g, f_1, ..., f_d) \colon f_i \in \mathcal{L}(F_i),\ i = 1, ..., d,\ h, g \in \mathcal{R} \text{ s.t. } h^- - g + \sum_{i=1}^{d} f_i \geq \varphi\right\}$$

and for each admissible function the corresponding value of the objective function is given by

$$\overline{Q}(h^-) - \underline{Q}(g) + \sum_{i=1}^{d} \mathbb{E}_i[f_i].$$

Again, for our reduction scheme, we consider a subclass of admissible pairs (h, g) such that $h, g \in \mathcal{R}^r$, i.e.

$$\overline{\mathcal{A}}^r := \left\{(h, g) \colon h, g \in \mathcal{R}^r \text{ s.t. } h^- - g \geq \varphi\right\}$$

We now turn to the formal construction of admissible functions in $\overline{\mathcal{A}}^r$ with an auxiliary version of the multiset inclusion-exclusion principle for intersections.

Lemma 4.8. *Let $B_1^n, ..., B_d^n \subset \mathbb{R}^d$ for $n = 1, ..., k$ and $k \in \mathbb{N}$ and define*

$$G_{(i_1,...,i_k)} := (B_{i_1}^1 \cap \cdots \cap B_{i_k}^k) \quad \textit{for } (i_1, ..., i_k) \in \{1, ..., d\}^k, \quad \textit{and} \quad \mathcal{B} = \bigcap_{n=1}^{k} \bigcup_{l=1}^{d} B_l^n.$$

Moreover, for an enumeration $\{\mathbf{i}^1, ..., \mathbf{i}^{dk}\}$ of the set $\{1, ..., d\}^k$ define the multisets

$$\langle \mathcal{G}^o, f^o \rangle, \quad \mathcal{G}^o := \{G_{\mathbf{i}^{n_1}} \cap \cdots \cap G_{\mathbf{i}^{n_m}} : 1 \leq n_1 < \cdots < n_m \leq k,\ m \textit{ odd}\} \tag{4.5}$$
$$\langle \mathcal{G}^e, f^e \rangle, \quad \mathcal{G}^e := \{G_{\mathbf{i}^{n_1}} \cap \cdots \cap G_{\mathbf{i}^{n_m}} : 1 \leq n_1 < \cdots < n_m \leq k,\ m \textit{ even}\} \tag{4.6}$$

where

$$f^o(G) = |\{(\mathbf{i}^{n_1}, ..., \mathbf{i}^{n_m}) \colon 0 \leq n_1 < \cdots < n_m \leq k,\ m \textit{ odd},\ G = G_{\mathbf{i}^{n_1}} \cap \cdots \cap G_{\mathbf{i}^{n_m}}\}|$$

and f^e is defined analogously. Then it holds that

$$\mathbb{1}_{\mathcal{B}} = \sum_{G \in \mathcal{G}^o} \left(f^o(G) - f^e(G)\right)^+ \mathbb{1}_G - \sum_{G \in \mathcal{G}^e} \left(f^e(G) - f^o(G)\right)^+ \mathbb{1}_G.$$

Proof. Since the union and the intersection of sets commute we have that $\mathcal{B} = \mathcal{G}_{\mathbf{i}^1} \cup \cdots \cup \mathcal{G}_{\mathbf{i}^{kd}}$ and hence the statement follows by a straight-forward application of Lemma 4.3. □

We are now ready to establish an explicit construction of admissible pairs $(h, g) \in \overline{\mathcal{A}}^r$. To this end, let us denote for $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathbb{R}^d$ and an enumeration $\{\mathbf{i}^1, ..., \mathbf{i}^{dk}\}$ of $\{1, ..., d\}^k$

$$\mathrm{U}_{\mathbf{i}^n} := \min(\mathrm{pr}_{i_1}(\mathbf{u}^1), ..., \mathrm{pr}_{i_k}(\mathbf{u}^k)) \quad \text{for } (i_1, ..., i_k) = \mathbf{i}^n, n = 1, ..., dk,$$

where $\mathrm{pr}_i(\mathbf{u}) := (\infty, ..., \infty, u_i, \infty, ..., \infty)$ for $i \in \{1, ..., d\}$. Moreover, we define

$$\begin{aligned}\mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k) &:= \{\min(\mathrm{U}_{\mathbf{i}^{n_1}}, ..., \mathrm{U}_{\mathbf{i}^{n_m}})\colon 1 \le n_1 < \cdots < n_m \le k \;\; m \text{ odd}\} \\ &\qquad = \mathcal{M}^o(\mathrm{U}_{\mathbf{i}^1}, ..., \mathrm{U}_{\mathbf{i}^{dk}}) \\ \mathcal{W}^e(\mathbf{u}^1, ..., \mathbf{u}^k) &:= \{\min(\mathrm{U}_{\mathbf{i}^{n_1}}, ..., \mathrm{U}_{\mathbf{i}^{n_m}})\colon 1 \le n_1 < \cdots < n_m \le k \;\; m \text{ even}\} \\ &\qquad = \mathcal{M}^e(\mathrm{U}_{\mathbf{i}^1}, ..., \mathrm{U}_{\mathbf{i}^{dk}}).\end{aligned} \tag{4.7}$$

Finally, we write $\mathbf{u} < \mathbf{v}$ for vectors $\mathbf{u}, \mathbf{v} \in \overline{\mathbb{R}}^d$ such that $u_i < v_i$ for $i = 1, ..., d$.

Lemma 4.9. *Let $k \in \mathbb{N}$ and $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}^c_\psi(s)$. Define the functions*

$$h_{\boldsymbol{\varepsilon}} := \sum_{\mathbf{u} \in \mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k)} \Lambda_{\mathbf{u}-\boldsymbol{\varepsilon}}; \quad g := \sum_{\mathbf{u} \in \mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k)} \Lambda_{\mathbf{u}},$$

for l^o and l^e being the multiplicity functions associated to $\mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k)$ and $\mathcal{W}^e(\mathbf{u}^1, ..., \mathbf{u}^k)$ respectively. Then for every $\boldsymbol{\varepsilon} > \mathbf{0}$ it holds that $(h_{\boldsymbol{\varepsilon}}, g)$ is admissible for the dual problem $\overline{D}_\varphi$, i.e. $(h_{\boldsymbol{\varepsilon}}, g) \in \overline{\mathcal{A}}^r$, and the value of the objective function is given by

$$\sum_{\mathbf{u} \in \mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \,\overline{Q}\big(F^-(\mathbf{u}+\boldsymbol{\varepsilon})\big) \; - \sum_{\mathbf{u} \in \mathcal{W}^e(\mathbf{u}^1, ..., \mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \,\underline{Q}\big(F(\mathbf{u})\big).$$

Proof. We need to show that $h^-_{\boldsymbol{\varepsilon}} - g \ge \varphi$. Note, that for every $(u_1, ..., u_d) \in \mathcal{U}^c_\psi(s)$ we have that

$$\big\{(x_1, ..., x_d) \in \mathbb{R}^d\colon x_1 \ge u_1, ..., x_d \ge u_d\big\}^c \supset \big\{(x_1, ..., x_d) \in \mathbb{R}^d\colon \psi(x_1, \cdots, x_d) < s\big\},$$

due to the fact that ψ is increasing in each coordinate. Hence, for $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}^c_\psi(s)$ and $B_n := \big\{(x_1, ..., x_d) \in \mathbb{R}^d\colon x_1 \ge u^n_1, ..., x_d \ge u^n_d\big\}$ for $n = 1, ..., k$ it follows that

$$\bigcap_{n=1}^k B^c_n \supset \big\{(x_1, ..., x_d) \in \mathbb{R}^d\colon \psi(x_1, \cdots, x_d) < s\big\}. \tag{4.8}$$

Moreover, for $n = 1, ..., k$ we have that

$$B^c_n = \bigcup_{i=1}^d \big\{(x_1, ..., x_d) \in \mathbb{R}^d\colon \mathbf{x} < \mathrm{pr}_i(u^n)\big\},$$

which follows from

$$\begin{aligned}\big\{(x_1, ..., x_d) \in \mathbb{R}^d\colon x_1 \ge u_1, ..., x_d \ge u_d\big\}^c &= \big([u_1, \infty) \times \cdots \times [u_d, \infty)\big)^c \\ &= \bigcup_{i=1}^d \mathbb{R} \times \cdots \times \mathbb{R} \times (-\infty, u_i) \times \mathbb{R} \times \cdots \times \mathbb{R} \\ &= \bigcup_{i=1}^d \big\{(x_1, ..., x_d) \in \mathbb{R}^d\colon \mathbf{x} < \mathrm{pr}(u^n)_i\big\}.\end{aligned}$$

Hence, we obtain

$$\bigcap_{n=1}^{k} B_n^c = \bigcap_{n=1}^{k} \bigcup_{i=1}^{d} \{(x_1, ..., x_d) \in \mathbb{R}^d \colon \mathbf{x} < \mathrm{pr}_i(u^n)\}. \tag{4.9}$$

Now, defining

$$H_i^n := \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon \mathbf{x} < \mathrm{pr}(u^n)_i)\big\}, \text{ for } i = 1, ..., d, n = 1, ..., k \text{ and}$$
$$\mathcal{G}_{(i_1,...,i_k)} := H_{i_1}^1 \cap \cdots \cap H_{i_k}^k,\ (i_1, ..., i_k) \in \{1, ..., d\}^k,$$

and applying Lemma 4.8 to equation (4.9) we arrive at

$$\mathbb{1}_{B_1^c \cap \cdots \cap B_k^c} = \sum_{G \in \mathcal{G}^o} \big(f^o(G) - f^e(G)\big)^+ \mathbb{1}_G - \sum_{G \in \mathcal{G}^e} \big(f^e(G) - f^o(G)\big)^+ \mathbb{1}_G,$$

where $\mathcal{G}^o, \mathcal{G}^e$ and f^o, f^e are defined in 4.5. Finally, note that for $1 \leq n_1 \leq \cdots \leq n_m \leq dk$

$$\mathcal{G}_{(i_1,...,i_k)} = \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon \mathbf{x} < \min\big(\mathrm{pr}_{i_1}(\mathbf{u}^1), ..., \mathrm{pr}_{i_k}(\mathbf{u}^k)\big)\big\},$$

so that with the definition of $\mathrm{U}_{\mathbf{i}}$ for $\mathbf{i} \in \{1, ..., d\}^k$ it follows for every $1 \leq n_1, ..., n_m \leq dk$ that

$$\begin{aligned}\bigcap_{l=1}^{m} \mathcal{G}_{\mathbf{i}^{n_l}} &= \bigcap_{l=1}^{m} \{(x_1, ..., x_d) \in \mathbb{R}^d \colon \mathbf{x} < \mathrm{U}_{\mathbf{i}^{n_l}}\} \\ &= \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon \mathbf{x} < \min(\mathrm{U}_{\mathbf{i}^{n_1}}, ..., \mathrm{U}_{\mathbf{i}^{n_m}})\big\},\end{aligned}$$

and thus $\mathbb{1}_{\mathcal{G}_{\mathbf{i}^{n_1}} \cap \cdots \cap \mathcal{G}_{\mathbf{i}^{n_m}}} = \Lambda^-_{\min(\mathrm{U}_{\mathbf{i}^{n_1}}, ..., \mathrm{U}_{\mathbf{i}^{n_m}})}$ for $\min(\mathrm{U}_{\mathbf{i}^{n_1}}, ..., \mathrm{U}_{\mathbf{i}^{n_m}}) \in \mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k) \cup \mathcal{W}^e(\mathbf{u}^1, ..., \mathbf{u}^k)$. In particular, using equation (4.8) and the fact that

$$l^o(\min(\mathrm{U}_{\mathbf{i}^{n_1}}, ..., \mathrm{U}_{\mathbf{i}^{n_m}})) = f^o(G)$$

for $G = \bigcap_{l=1}^m G_{\mathbf{i}^{n_l}} \in \mathcal{G}^o$ and vice versa for l^e, we conclude that

$$h_{\boldsymbol{\varepsilon}}^-(\mathbf{x}) - g(\mathbf{x}) \geq h_{\mathbf{0}}^-(\mathbf{x}) - g^-(\mathbf{x}) = \mathbb{1}_{B_1^c \cap \cdots \cap B_k^c}(\mathbf{x}) \geq \mathbb{1}_{\psi(\mathbf{x}) < s} = \varphi(\mathbf{x})$$

for all $\mathbf{x} \in \mathbb{R}^d$, which completes the proof. □

We are now in the position to establish our reduction scheme based on $\overline{D}_\varphi$. The proof of the following theorem is analogous to the proof of Theorem 4.6 and therefore omitted.

Theorem 4.10. *Let* $\varphi(x_1, ..., x_d) = \mathbb{1}_{\psi(x_1,...,x_d) < s}$ *where* $\psi \colon \mathbb{R}^d \to \mathbb{R}$ *is increasing in each coordinate and let*

$$\begin{aligned}\overline{D}_\varphi(k) := \inf\Bigg\{ &\sum_{\mathbf{u} \in \mathcal{W}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \overline{Q}\big(F(\mathbf{u})\big) \\ &- \sum_{\mathbf{u} \in \mathcal{W}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \underline{Q}\big(F(\mathbf{u})\big) \colon \mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_\psi^c(s)\Bigg\},\end{aligned} \tag{4.10}$$

where l^o and l^e are the canonical multiplicity functions associated to $\mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k)$ and $\mathcal{W}^e(\mathbf{u}^1, ..., \mathbf{u}^k)$ respectively. Then

$$\overline{D}_\varphi(k) \geq \overline{D}_\varphi(k+1) \geq ... \geq \overline{D}_\varphi.$$

Remark 4.11. Note that since $\varphi(x_1, ..., x_d) = \mathbb{1}_{\psi(x_1,...,x_d)<s}$ is not upper semicontinuous, strong duality as in Theorem 3.2 is not guaranteed. Nevertheless, $\overline{D}_\varphi(k)$ is an upper bound on the generalized Fréchet functional $\overline{P}_\varphi$. ♦

4.3. Sharp asymptotic bounds in the certainty limit

In general, the schemes $\underline{D}_\varphi(k)$ and $\overline{D}_\varphi(k)$ do not approximate the dual bounds $\underline{D}_\varphi$ and $\overline{D}_\varphi$ respectively for $k \to \infty$. In the homogeneous, complete dependence uncertainty case, i.e. $F_1 = \cdots = F_d = F$ and $\underline{Q} = W_d$ and $\overline{Q} = M_d$, Puccetti and Rüschendorf (2013) derived an explicit solution to the dual $\underline{D}_\varphi$ under additional requirements on the marginals. They showed, that the optimizer is of the form $d \cdot f$ for a piecewise linear function $f \in \mathcal{L}(F)$ which cannot be represented by the linear combinations in $\mathcal{R}$.

The counterpart to the situation of complete dependence uncertainty is the case of certainty, i.e. the limit when $\underline{Q}$ and $\overline{Q}$ converge from below and above respectively to a copula C. A natural feature of any bound on the expectation of φ using the information from $\underline{Q}$ and $\overline{Q}$ should be that it converges to $\mathbb{E}_C[\varphi]$ as $\underline{Q}, \overline{Q} \to C$. The following theorem shows that for $k \to \infty$ the reduced bounds $\underline{D}_\varphi(k)$ and $\overline{D}_\varphi(k)$ indeed converge to the desired object in the certainty limit. In order to make the dependence on the copula bounds explicit, we define for $k \in \mathbb{N}$ the functions $[\underline{D}_\varphi(k)](\underline{Q}, \overline{Q}) := \underline{D}_\varphi(k)$ and $[\overline{D}_\varphi(k)](\underline{Q}, \overline{Q}) := \overline{D}_\varphi(k)$ where $\underline{D}_\varphi(k), \overline{D}_\varphi(k)$ depend on the lower bound $\underline{Q}$ and the upper bound $\overline{Q}$.

Theorem 4.12. *Let $\varphi(x_1, ..., x_d) = \mathbb{1}_{\psi(x_1,...,x_d)<s}$ for $\psi \colon \mathbb{R}^d \to \mathbb{R}$ increasing in each coordinate. Then it holds for any copula C and sequences of quasi-copulas $(\underline{Q}^j)_{j=1,2...}$ and $(\overline{Q}^j)_{j=1,2...}$ with $\underline{Q}^j \leq C \leq \overline{Q}^j$ for all $j \in \mathbb{N}$ and $\underline{Q}^j, \overline{Q}^j \to_j C$ pointwise, that*

$$\lim_j \sup_k [\underline{D}_\varphi(k)](\underline{Q}^j, \overline{Q}^j) = \mathbb{P}_C(\psi(X_1, ..., X_d) < s),$$

when ψ is upper semicontinuous and

$$\lim_j \inf_k [\overline{D}_\varphi(k)](\underline{Q}^j, \overline{Q}^j) = \mathbb{P}_C(\psi(X_1, ..., X_d) < s),$$

when ψ is lower semicontinuous.

Proof. We show that

$$\lim_j \sup_k [\underline{D}_\varphi(k)](\underline{Q}^j, \overline{Q}^j) = \mathbb{P}_C(\psi(X_1, ..., X_d) < s).$$

The proof for $\overline{D}_\varphi(k)$ follows along similar lines.

First, note that for each $\varepsilon > 0$ there exists a sequence $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_\psi(s)$, such that for

$$h_k := \sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \left(l^o(\mathbf{u}) - l^e(\mathbf{u})\right)^+ \Lambda_{\mathbf{u}}; \quad g_k^- := \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)} \left(l^e(\mathbf{u}) - l^o(\mathbf{u})\right)^+ \Lambda_{\mathbf{u}}^-,$$

where $\mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k)$ and $\mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^k)$ are the multisets defined in (4.2), we have that

$$\left|\mathbb{P}_C(\psi(X_1, ..., X_d) < s) - \mathbb{E}_C\left[h_k - g_k^-\right]\right| < \frac{\varepsilon}{2} \tag{4.11}$$

with $\mathbb{P}_C(\psi(X_1, ..., X_d) < s) = \mathbb{E}_C[\varphi]$. To verify the existence of such a sequence one can choose $m \in \mathbb{N}$ such that $\mathbb{P}_C\big(\mathcal{U}_\psi(s)\big) - \mathbb{P}_C\big(\mathcal{U}_\psi(s) \cap [-m, m]^d\big) < \frac{\varepsilon}{4}$ and $(\mathbf{u}^n)_{n=1,...,k}$ can be any discretization of the set $\mathcal{U}_\psi(s) \cap [-m, m]^d$, whose mesh converges to zero for $k \to \infty$. Since ψ is upper semicontinuous, $\mathcal{U}_\psi(s)$ is an open set and we can chose k such that

$$\left|\mathbb{P}_C\big(\mathcal{U}_\psi(s) \cap [-m, m]^d\big) - \mathbb{E}_C\left[h_k - g_k^-\right]\right| < \frac{\varepsilon}{4}.$$

It follows that

$$\begin{aligned}
&\left|\mathbb{P}_C(\psi(X_1, ..., X_d) < s) - \mathbb{P}_C\big(\mathcal{U}_\psi(s) \cap [-m, m]^d\big) + \mathbb{P}_C\big(\mathcal{U}_\psi(s) \cap [-m, m]^d\big) - \mathbb{E}_C\left[h_k - g_k^-\right]\right| \\
&\leq \left|\mathbb{P}_C(\psi(X_1, ..., X_d) < s) - \mathbb{P}_C\big(\mathcal{U}_\psi(s) \cap [-m, m]^d\big)\right| \\
&\qquad + \left|\mathbb{P}_C\big(\mathcal{U}_\psi(s) \cap [-m, m]^d\big) - \mathbb{E}_C\left[h_k - g_k^-\right]\right| < \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2}.
\end{aligned}$$

Moreover, from the fact that $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_\psi(s)$ and the proof of Lemma 4.5 it follows that $h_k - g_k^- \leq \varphi$ for all $k \in \mathbb{N}$, and the corresponding value of the objective function is given by

$$\sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \left(l^o(\mathbf{u}) - l^e(\mathbf{u})\right)^+ \underline{Q}\big(F(\mathbf{u})\big) - \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)} \left(l^e(\mathbf{u}) - l^o(\mathbf{u})\right)^+ \overline{Q}\big(F^-(\mathbf{u})\big) \leq \left[\underline{D}_\varphi(k)\right](\underline{Q}^j, \overline{Q}^j). \tag{4.12}$$

We proceed by showing that the convergence

$$\lim_j \sup_k \left[\underline{D}_\varphi(k)\right](\underline{Q}^j, \overline{Q}^j) = \mathbb{P}_C(\psi(X_1, ..., X_d) < s)$$

holds. To this end, fix an arbitrary $\varepsilon > 0$ and a corresponding sequence $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_\psi(s)$ as in (4.11). Moreover, the fact that quasi-copulas are Lipschitz continuous yields, by an application of the Arzelà-Ascoli Theorem, that $\underline{Q}^j \to_j C$ and $\overline{Q}^j \to_j C$ uniformly. Thus, for

$$p := \sum_{\mathbf{u}\in\mathcal{M}^o\mathbf{u}^1,...,\mathbf{u}^k)} l^o(\mathbf{u}) + \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)} l^e(\mathbf{u})$$

we can choose an $j \in \mathbb{N}$ so that

$$\|C - \underline{Q}^j\|_\infty + \|C - \overline{Q}^j\|_\infty < \frac{\varepsilon}{2p}. \tag{4.13}$$

With this choice of k and j we arrive at

$$\begin{aligned}
&\left|\mathbb{P}_C(\psi(X_1, ..., X_d) < s) - \left[\underline{D}_\varphi(k)\right](\underline{Q}^j, \overline{Q}^j)\right| \\
&\leq \left|\mathbb{P}_C(\psi(X_1, ..., X_d) < s) - \mathbb{E}_C\left[h_k - g_k^-\right] + \mathbb{E}_C\left[h_k - g_k^-\right] - \left[\underline{D}_\varphi(k)\right](\underline{Q}^j, \overline{Q}^j)\right| \\
&\leq \left|\mathbb{P}_C(\psi(X_1, ..., X_d) < s) - \mathbb{E}_C\left[h_k - g_k^-\right]\right| \\
&+ \Big|\mathbb{E}_C\left[h_k - g_k^-\right] - \Big(\sum_{\mathbf{u} \in \mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k)} \left(l^o(\mathbf{u}) - l^e(\mathbf{u})\right)^+ \underline{Q}^j\left(F(\mathbf{u})\right) \\
&\qquad - \sum_{\mathbf{u} \in \mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^k)} \left(l^e(\mathbf{u}) - l^o(\mathbf{u})\right)^+ \overline{Q}^j\left(F^-(\mathbf{u})\right)\Big)\Big| \\
&< \frac{\varepsilon}{2} + \Big| \sum_{\mathbf{u} \in \mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k)} \left(l^o(\mathbf{u}) - l^e(\mathbf{u})\right)^+ \Big(C(F(\mathbf{u})) - \underline{Q}^j(F(\mathbf{u}))\Big) \\
&\qquad - \sum_{\mathbf{u} \in \mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^k)} \left(l^o(\mathbf{u}) - l^e(\mathbf{u})\right)^+ \Big(C(F^-(\mathbf{u})) - \overline{Q}^j(F^-(\mathbf{u}))\Big)\Big| \\
&< \frac{\varepsilon}{2} + p\big(\|C - \underline{Q}^j\|_\infty + \|C - \overline{Q}^j\|_\infty\big) < \frac{\varepsilon}{2} + p\frac{\varepsilon}{2p} = \varepsilon.
\end{aligned}$$

The second inequality is a consequence of equation (4.12) and the fact that $\mathbb{E}_C[h_k - g_k^-] \geq \left[\underline{D}_\varphi(k)\right](\underline{Q}^j, \overline{Q}^j)$. The third inequality follows from equation (4.11) and last inequality holds due to equation (4.13).

Finally, since ε was arbitrary we have shown that

$$\lim_j \sup_k \left[\underline{D}_\varphi(k)\right](\underline{Q}^j, \overline{Q}^j) = \mathbb{P}_C(\psi(X_1, ..., X_d) < s)$$

holds and hence the proof is complete. □

5. Illustrations and numerical examples

In this section we provide an informal description of the reduction schemes in Sections 4.1 and 4.2 in order to illustrate the underlying idea. Furthermore, we provide several numerical examples comparing the performance of our reduction scheme to the improved standard bounds given e.g. in Embrechts et al. (2003).

A graphical illustration of $\underline{D}_\varphi(k)$

For a graphical illustration of the scheme $\underline{D}_\varphi(k)$ let us consider $\psi(x_1, x_2) = x_1 + x_2$ and F_1, F_2 uniform distributions on $[0, 1]$. Due to assumption **(A1)** and **(A2)**, we consider admissible functions which are sums of indicator functions of rectangular regions in $\mathcal{U}_\psi(s)$, as in

$$h - g = \sum_{n=1}^{k} \Lambda_{\mathbf{u}^k} - \sum_{n=1}^{m} \Lambda_{\mathbf{v}^n} \leq \mathbb{1}_{x_1+x_2<s}$$

for $\mathbf{u}^1, ..., \mathbf{u}^k, \mathbf{v}^1, ..., \mathbf{v}^m \in \mathcal{U}_\psi(s)$ and $k, m \in \mathbb{N}$.The corresponding value of the objective function to be maximized is given by $\underline{Q}(h) - \overline{Q}(g^-)$ for each pair $h, g \in \mathcal{R}^r$. Note that due to the continuity of the marginal distributions it holds that $\overline{Q}(g^-) = \overline{Q}(g)$. Figure 1 illustrates the structure of admissible functions of this type that we shall consider.

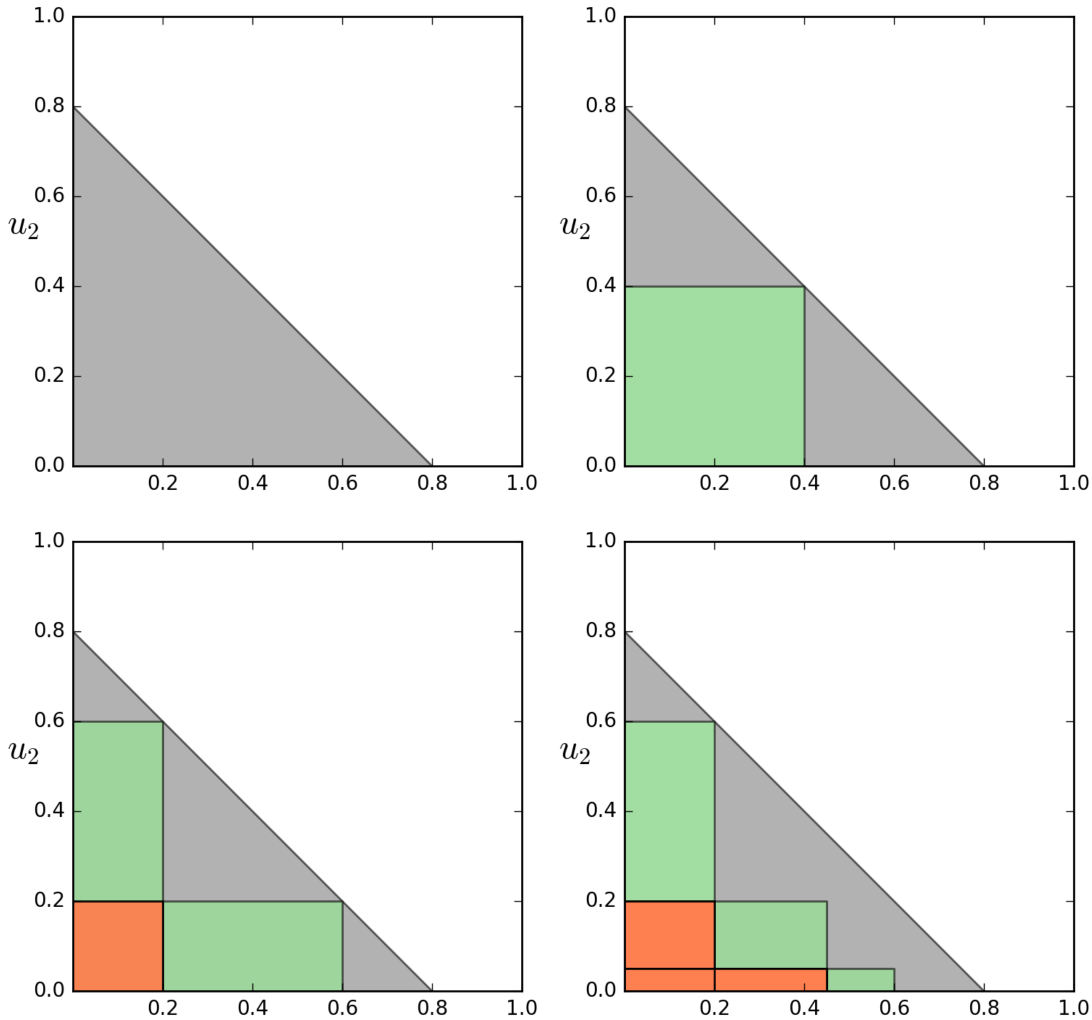

Figure 1: Constrained set of admissible functions

The gray triangular region in Figure 1, corresponds to the area

$$\mathcal{U}_\psi(s) = \{(u_1, u_2) \in [0,1]^2 \colon u_1 + u_2 < s\}$$

for $s = 0.8$. The upper LHS simply depicts the region $\mathcal{U}_\psi(s)$. The green area in the upper RHS figure corresponds to the rectangle $[0, 0.4]^2 \subset \mathcal{U}_\psi(s)$ as induced by the function $h = \Lambda_{(0.4,0.4)}$, i.e. $\Lambda_{(0.4,0.4)}(u_1, u_2) = \varphi(u_1, u_2) = 1$ for all $(u_1, u_2) \in [0, 0.4]^2$. The value of the objective function h is given by

$$\underline{Q}(\Lambda_{(0.4,0.4)}) = \underline{Q}(F_1(0.4), F_2(0.4)) = \underline{Q}(0.4, 0.4).$$

Similarly, the lower LHS represents the rectangles $[0, 0.2] \times [0, 0.6]$ and $[0, 0.6] \times [0, 0.2]$ induced by $\Lambda_{(0.2,0.6)}$ and $\Lambda_{(0.6,0.2)}$. The red area corresponds to $[0, 0.2] \times [0, 0.2]$ where an overlap occurs due to

$$h(\mathbf{u}) = \Lambda_{(0.2,0.6)}(\mathbf{u}) + \Lambda_{(0.6,0.2)}(\mathbf{u}) = 2 > \varphi(\mathbf{u}) \quad \text{for all } \mathbf{u} \in [0, 0.2] \times [0, 0.2].$$

This overlap is then compensated by applying the inclusion-exclusion principle and subtracting $g = \Lambda_{(0.2,0.2)}$, yielding the admissible function

$$h - g = \Lambda_{(0.2,0.6)} + \Lambda_{(0.6,0.2)} - \Lambda_{(0.2,0.2)}.$$

The respective value of the objective function is equal to $\underline{Q}(0.2, 0.6) + \underline{Q}(0.6, 0.2) - \overline{Q}(0.2, 0.2)$. Finally, the lower RHS represents the function constructed by

$$h = \Lambda_{(0.2,0.6)} + \Lambda_{(0.45,0.2)} + \Lambda_{(0.6,0.05)}$$

and an appropriate compensation of the overlap by $g = \Lambda_{(0.2,0.2)} + \Lambda_{(0.45,0.05)}$ so that $(h, g) \in \underline{\mathcal{A}}$ and the corresponding value of the objective function is equal to

$$\underline{Q}(0.2, 0.6) + \underline{Q}(0.45, 0.2) + \underline{Q}(0.6, 0.05) - \overline{Q}(0.2, 0.2) - \overline{Q}(0.45, 0.05).$$

Note, that the construction of (h, g) depends entirely on the choice $\mathbf{u}_1, ..., \mathbf{u}_k \in \mathbb{R}^d$. In particular, maximizing over all (h, g) that are constructed in this way amounts to an optimization over $\mathbf{u}_1, ..., \mathbf{u}_k \in \mathcal{U}_\psi(s)$.

Lemma 4.5 formalizes this construction of admissible (h, g) whereas Theorem 4.6 establishes the optimization over such (h, g) yielding bounds on $\underline{D}_\varphi$.

A graphical illustration of $\overline{D}_\varphi(k)$

Using again $\psi(x_1, x_2) = x_1 + x_2$ and F_1, F_2 uniform distributions on $[0, 1]$ let us illustrate the idea of the scheme $\overline{D}_\varphi(k)$. This time, (h, g) with $h, g \in \mathcal{R}^r$ are admissible when

$$h - g = \sum_{n=1}^{k} \Lambda_{\mathbf{u}^k} - \sum_{n=1}^{m} \Lambda_{\mathbf{v}^n} \geq \mathbb{1}_{x_1 + x_2 < s}.$$

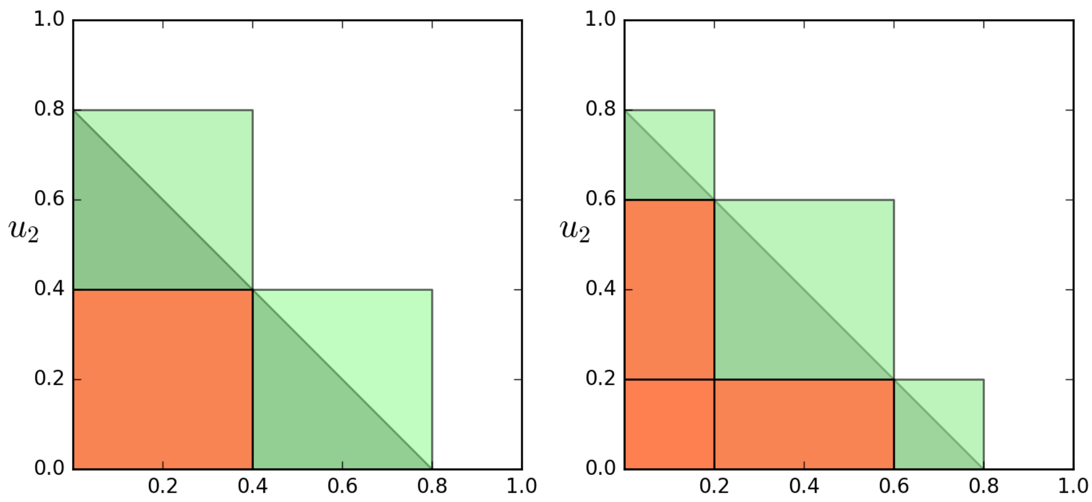

Figure 2: Constrained set of admissible functions

Figure 1 illustrates two possible constructions of admissible pairs (h, g). Again the green area corresponds to $\{\mathbf{x} \in [0,1]^2 \colon h(\mathbf{x}) = 1\}$ whereas the red shaded area marks overlaps ($h(\mathbf{x}) > 1$) which we compensate using the inclusion exclusion principle. The LHS corresponds to

$$h - g = \Lambda_{(0.8,0.4)} + \Lambda_{(0.4,0.8)} - \Lambda_{(0.4,0.4)} \in \overline{\mathcal{A}}$$

and the respective value of the objective function amounts to $\overline{Q}(0.8, 0.4)+\overline{Q}(0.4, 0.8)-\underline{Q}(0.4, 0.4)$. The RHS represents the admissible function given by $h - g$ for

$$h = \Lambda_{(0.8,0.2)} + \Lambda_{(0.2,0.8)} + \Lambda_{(0.6,0.6)}; \quad g = \Lambda_{(0.2,0.6)} + \Lambda_{(0.6,0.2)}$$

with corresponding value of the objective function

$$\overline{Q}(0.8, 0.2) + \overline{Q}(0.2, 0.8) - \overline{Q}(0.6, 0.6) - \underline{Q}(0.2, 0.6) - \underline{Q}(0.6, 0.2).$$

Note, that in contrast to $\underline{D}_\varphi(k)$ it does not suffice to consider $h = \sum_{n=1}^{k} \mathbf{u}^k$ for $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_\psi(s)$. A construction of admissible functions in the spirit of Section 4.1 is however possible if we formulate it in terms of indicator functions of upper level sets of the form

$$\{(x_1, x_2) \colon x_1 \geq u_1,\ x_2 \geq u_2\},$$

for $\mathbf{u}$ in the complement $\mathcal{U}^c_\psi(s)$. Returning to the LHS of Figure 2 we then note that the region

$$\{\mathbf{x} \in \mathbb{R}^2 \colon \Lambda_{(0.8,0.4)}(\mathbf{x}) + \Lambda_{(0.4,0.8)}(\mathbf{x}) - \Lambda_{(0.4,0.4)}(\mathbf{x}) = 1\}$$

can be expressed in terms of complements of upper level sets via

$$\big([0.4, 1] \times [0.4, 1]\big)^c \cap \big([0, 1] \times [0.8, 1]\big)^c \cap \big([0.8, 1] \times [0, 1]\big)^c,$$

and a similar representation holds for the RHS. Moreover, we can represent the complements via

$$\big([u_1,1]\times[u_2,1]\big)^c = ([0,u_1]\times[0,1])\cup([0,1]\times[0,u_2]),$$

where the right-hand side of the equation is the union of sets that can be evaluated by means of the quasi-copulas $\underline{Q},\overline{Q}$. This construction is made rigorous in Lemma 4.9 and the resulting optimization is provided in Theorem 4.10.

The following numerical examples show, how the reduction schemes can be applied in order to account for copula information in the computation of VaR estimates. Our results are compared to the improved standard bounds using the same information.

Example 5.1. Consider an $\mathbb{R}^5$-valued risk vector $(X_1,\dots,X_5)$ with copula C and Pareto$_2$-marginals. We assume that the copula C lies in the vicinity of a reference copula C^* as measured by the Kolmogorov–Smirnov distance, i.e.

$$\|C-C^*\|_\infty \le \delta$$

for some $\delta>0$. This situation typically occurs when estimating copulas from empirical data. In this case, one can think of C^* as being the empirical copula obtained from the available data, while C is an estimator from a parametric family of copulas so as to minimize some distance to C^*. In this case δ corresponds the residual error resulting from the estimation procedure.

Theorem 5.4 in Lux and Papapantoleon (2016) establishes pointwise upper and lower bounds on the set of copulas in the δ-neighborhood of C^*, for a large class of distances satisfying some monotonicity requirements. In particular, using the explicit representation of these bounds given in Lux and Papapantoleon (2016, Lemma 5.7) we obtain that

$$\begin{aligned}\underline{Q}^{\|\cdot\|_\infty,\delta}(\mathbf{u}) := \max\{C^*(\mathbf{u})-\delta, W_5(\mathbf{u})\} &\le C(\mathbf{u})\\ &\le \min\{C^*(\mathbf{u})+\delta, M_5(\mathbf{u})\} =: \overline{Q}^{\|\cdot\|_\infty,\delta}(\mathbf{u}),\end{aligned} \tag{5.1}$$

for all $\mathbf{u}\in\mathbb{I}^5$. Furthermore, in order to compare our results to the improved standard bounds on VaR, we assume that the survival copula $\widehat{C}$ is such that

$$\|\widehat{C}-\widehat{C}^*\|_\infty \le \delta.$$

Hence, from Lux and Papapantoleon (2016, Corollary 5.8) we obtain similar bounds $\widehat{\underline{Q}}^{\|\cdot\|_\infty,\delta}$ and $\widehat{\overline{Q}}^{\|\cdot\|_\infty,\delta}$ on the survival copula of $(X_1,\dots,X_5)$ such that

$$\widehat{\underline{Q}}^{\|\cdot\|_\infty,\delta} \le \widehat{C} \le \widehat{\overline{Q}}^{\|\cdot\|_\infty,\delta}. \tag{5.2}$$

We then translate the bounds in (5.1) and the corresponding bounds on the survival copula into estimates on the VaR of the sum $X_1+\cdots+X_5$ by means the reduction schemes presented in Section 4 and Appendix A. These estimates are compared to the improved standard bounds, which are given by the inverses of the following bounds on the distribution function of $X_1 + \cdots + X_5$, along the variable s:

$$\mathbb{P}(X_1 + \cdots + X_5 < s) \geq \sup_{x_1+\cdots+x_5<s} \underline{Q}^{\|\cdot\|_\infty,\delta}\big(F_1(x_1), ..., F_5(x_5)\big)$$
$$\mathbb{P}(X_1 + \cdots + X_5 < s) \leq \inf_{x_1+\cdots+x_5\geq s} 1 - \underline{\widehat{Q}}^{\|\cdot\|_\infty,\delta}\big(F_1(x_1), ..., F_5(x_5)\big).$$

The following tables show the improved standard bounds using $\underline{Q}^{\|\cdot\|_\infty,\delta}$ and $\underline{\widehat{Q}}^{\|\cdot\|_\infty,\delta}$ for different levels of the confidence threshold α, as well as the bounds obtained by inverting $\underline{D}_\varphi(k), \underline{\widehat{D}}_\varphi(k)$ and $\overline{D}_\varphi(k), \underline{\widehat{D}}_\varphi(k)$, for $k = 6$ and $\varphi(x_1, ..., x_5) = \mathbb{1}_{x_1+\cdots x_5<s}$, along the variable s. We thus obtain two lower and two upper VaR estimates via the reduction schemes for each α, of which the largest lower bound and the lowest upper bound respectively are reported in each of the tables. For the computation we assume that the reference copula C^* is Gaussian with equicorrelation matrix and correlation coefficient ρ. The bounds consistently improve for increasing $k = 1, ..., 6$ while for $k \geq 7$ no further improvement was obtained. For the sake of legibility, the results are rounded to one decimal place. Table 1 shows the VaR estimates for different levels of the correlation of the reference copula and $\delta = 0.0001$ and different thresholds for the distance δ. Note, that in order to derive informative bounds on the tail of the distribution of the sum the seemingly small choice of the threshold δ is appropriate.

	$\delta = 0.0001$								
	$\rho = -0.1$			$\rho = 0.4$			$\rho = 0.8$		
α	i. standard (low : up)	scheme (low : up)	impr. %	i. standard (low : up)	scheme (low : up)	impr. %	i. standard (low : up)	scheme (low : up)	impr. %
0.95	3.5 : 44.8	8.8 : 24.0	63%	3.7 : 41.2	8.0 : 26.7	50%	7.8 : 31.3	9.6 : 24.8	35%
0.99	9.0 : 106.3	19.9 : 44.8	74%	9.1 : 102.7	19.0 : 62.5	54%	17.8 : 82.0	21.5 : 64.5	33%
0.995	13.3 : 152.1	27.0 : 60.8	76%	13.3 : 149.4	25.8 : 90.5	52%	24.3 : 119.0	28.5 : 91.6	33%
	$\delta = 0.0005$								
0.95	3.4 : 45.0	8.2 : 24.8	60%	3.6 : 41.2	7.2 : 28.1	44%	7.8 : 31.4	9.2 : 26.2	28%
0.99	9.0 : 106.2	15.9 : 56.7	58%	9.0 : 105.3	14.9 : 80.8	32%	17.4 : 84.9	18.6 : 82.2	6%
0.995	13.3 : 153.0	19.0 : 90.0	49%	13.3 : 153.0	18.0 : 153.0	3%	23.4 : 126.0	22.8 : 125.0	0%

Table 1: Improved standard bounds on VaR of $X_1 + \cdots + X_5$ and VaR estimates via reduction schemes for different correlation parameters ρ and distance thresholds δ.

For $\delta = 0.0001$, the improvement obtained by including two-sided information via the reduction scheme ranges from 33% in the case of high positive correlation to 76% in the case of weak negative correlation. Overall, the improvement is more pronounced when weak or negative correlation is prescribed. Note, that in all cases the bounds improve the sharp unconstrained bounds, i.e. the sharp VaR bounds when no dependence information at all is available, which can be approximated by means of the Rearrangement Algorithm; see Embrechts et al. (2013). The unconstrained bounds amount to 3.2–39.6 for $\alpha = 0.95$, 4–94.8 when $\alpha = 0.99$ and 4.3–136 for $\alpha = 0.995$. For $\delta = 0.0005$, the improvement is – as expected – considerably weaker, especially when the correlation or the confidence threshold α are high and no improvement over the standard bounds is obtained for $\rho = 0.8$ and $\alpha = 0.995$. $\diamondsuit$

Remark 5.2. Example 5.1 illustrates how the methods developed in Section 4 can be used to evaluate model risk in ways that comply with the five fundamental criteria for coherent risk aggregation stated in Cambou and Filipović (2015). For instance, an institution may consider the reference copula C^* or the corresponding survival copula as internal dependence models for their risk exposures and penalties apply when the internal model yields VaR estimates that are too far from the robust bounds obtained via the reduction schemes. Then it holds that our method

1. leads to lower penalties for conservative models, i.e. low penalty is applied to a reference copula (internal model) that yields an aggregated VaR close to the robust (worst-case) estimate.
2. By choosing a suitable distance (not necessarily Komogorov–Smirnov) focus can be placed on tail dependence. Note, that Lux and Papapantoleon (2016, Lemma 5.7) provides improved Fréchet–Hoeffding bounds for a large class of distances.
3. The distance to the internal model can be controlled easily.
4. Capital requirements derived from the robust VaR bounds are in turn robust to model risk as long as the underlying assumptions are met.
5. The bounds are tractable and numerically computable. $\blacklozenge$

Example 5.3. We now consider an $\mathbb{R}^4$-valued risk vector (X_1, X_2, X_3, X_4) with copula C and Pareto$_2$-marginals. Moreover, we assume that

$$C^{\underline{\Sigma}} \leq C \leq C^{\overline{\Sigma}}$$

where $C^{\underline{\Sigma}}$ and $C^{\overline{\Sigma}}$ denote 4-dimensional Gaussian copulas with correlation matrices $\underline{\Sigma} = (\underline{\rho}_{ij})_{i,j=1,\ldots,4}$ and $\overline{\Sigma} = (\overline{\rho}_{ij})_{i,j=1,\ldots,4}$ respectively. Also, we assume that $\underline{\rho}_{ij} \leq \overline{\rho}_{ij}$ for $i, j = 1, \ldots, 4$, which by Slepian's Lemma guarantees non-degeneracy in the sense that $C^{\underline{\Sigma}} \leq C^{\overline{\Sigma}}$; c.f. Gupta, Eaton, Olkin, Perlman, Savage, and Sobel (1972, Theorem 5.1).

This corresponds to a situation of correlation uncertainty which occurs naturally in applications. Whenever correlation is estimated from data one obtains, rather than an exact estimate, a confidence interval for the pairwise correlations $(\underline{\rho}_{ij}, \overline{\rho}_{ij}) \subset [-1, 1]$, in which the parameters lie with high probability. Moreover, we assume that bounds on the survival function $\widehat{C}$ are given by the respective survival functions of $C^{\underline{\Sigma}}$ and $C^{\overline{\Sigma}}$, i.e.

$$C^{\underline{\Sigma}}(\mathbf{1} - \cdot) \leq \widehat{C} \leq C^{\overline{\Sigma}}(\mathbf{1} - \cdot).$$

We then relate the bounds on C and $\widehat{C}$ respectively to the VaR of $X_1 + \cdots + X_4$, using our reduction schemes and again we compare the results to the improved standard bounds obtained from $C^{\underline{\Sigma}}$ and $C^{\underline{\Sigma}}(\mathbf{1} - \cdot)$. Table 2 shows the results for different confidence levels α, assuming that $\underline{\Sigma}$ and $\overline{\Sigma}$ are equicorrelation matrices with correlation parameters $\underline{\rho}$ and $\overline{\rho}$ respectively. The VaR estimates were obtained by inverting $\underline{D}_\varphi(k)$ and $\overline{D}_\varphi(k)$ as well as $\widehat{\underline{D}}_\varphi(k)$ and $\widehat{\overline{D}}_\varphi(k)$ for $\varphi(x_1, ..., x_4) = \mathbb{1}_{x_1+\dots+x_4<s}$ and $k = 5$ along the variable s. Thus, we obtain two upper and two lower VaR estimates of which the largest lower bound and the lowest upper bound are reported. The bounds consistently improve for increasing $k = 1, ..., 5$ while for $k \geq 6$ no further improvement was obtained.. For the sake of legibility the results are rounded to full integers.

	$\underline{\rho} = -0.1, \overline{\rho} = 0.2$			$\underline{\rho} = 0.3, \overline{\rho} = 0.5$		
α	i. standard (low : up)	scheme (low : up)	impr. %	i. standard (low : up)	scheme (low : up)	impr. %
0.95	3 : 32	8 : 26	38	1 : 30	7 : 29	24
0.99	9 : 74	20 : 52	51	2 : 74	18 : 63	37
0.995	13 : 104	26 : 70	52	3 : 104	25 : 86	40

Table 2: Improved standard bounds on VaR of $X_1 + \cdots + X_4$ and VaR estimates computed via reduction schemes using $C^{\underline{\Sigma}}$ and $C^{\overline{\Sigma}}$.

The improvement of the spread ranges from 24% in the case of moderate positive correlation up to 52% in the case of low correlation. Moreover, the improvement is particularly pronounced for high levels of the confidence threshold α. Moreover, all bounds are narrower than the unconstrained VaR bounds without dependence information which amount to 3–31 for $\alpha = 0.95$, 3–74 for $\alpha = 0.99$ and 3–109 when $\alpha = 0.995$. ◊

Example 5.4. In this example we consider a situation where the joint laws of lower dimensional subgroups of the risk vector are available. This relates to the results and applications considered in Bignozzi et al. (2015), Puccetti, Rüschendorf, Small, and Vanduffel (2017) and Rüschendorf and Witting (2017). Specifically, we consider a 16-dimensional homogeneous risk vector $\mathbf{X} = (X_1, ..., X_{16})$ having log-normally distributed marginals $F = F_1, ..., F_{16}$ with mean $\mu = 0$ and standard deviation $\sigma = 1$. We decompose the risk vector into m smaller subgroups by means of a partition $I_1, ..., I_m$ of the index set $\{1, ..., 16\}$, i.e. $\bigcup_{j=1}^m I_j = \{1, ..., 16\}$. Define the subvectors of $\mathbf{X}$ by $X_{I_j} := (X_i)_{i\in I_j}$ for $j = 1, ..., m$. We assume that the joint law of each subvector is given. Similar to Bignozzi et al. (2015), we suppose that the subgroups are comonotonic so that their copula is equal to the upper Fréchet-Hoeffding bound, hence

$$\mathbb{P}\Big(\bigcap_{i\in I_j}\{X_i \leq x_i\}\Big) = \min_{i\in I_j}\{F(x_i)\}.$$

Due to the homogeneity of the marginals and the assumption of comonotonicity within the groups, we obtain an explicit distribution for the sum $\sum_{i\in I_j} X_i =: X'_{I_j}$, i.e.

$$\mathbb{P}\Big(\sum_{i\in I_j} X_i \leq x\Big) = F\Big(\frac{x}{|I_j|}\Big) =: F'_j(x) \quad \text{for all } x \in \mathbb{R},\ j = 1, ..., m.$$

Moreover, we assume that model ambiguity is prevalent in form of an unknown copula between the individual subgroups. Along with the above argument this is equivalent to the copula C_m of $(X'_{I_1}, ..., X'_{I_m})$ being unknown. Specifically, we assume that C_m and the corresponding survival copula $\widehat{C}_m$ lie between two Frank copulas, i.e.

$$C^{\theta_1} \le C_m \le C^{\theta_2} \quad \text{and} \quad \widehat{C}^{\theta_1} \le \widehat{C}_m \le \widehat{C}^{\theta_2},$$

for $\theta_1 < \theta_2$, where C^θ denotes the Frank copula with parameter θ. Recall that the Frank copula is positively ordered w.r.t. θ and thus $C^{\theta_1} \le C^{\theta_2}$ which guarantees that the problem is not degenerate.

In practice, this setting is suitable to aggregate e.g. the total risk of a financial institution consisting of multiple branches, some of which are strongly positively dependent while the dependence between several groups of branches is unknown; for details see also Bignozzi et al. (2015).

We then analyse the effect of different groups sizes on the improved VaR bounds for the aggregation $X_1 + \cdots + X_{16}$. To this end, we consider three different decompositions, each consisting of a partition of the index set $\{1, ..., 16\}$ into groups of equal size with $m = 2, 4, 8$ constituents. We then use the available information on the subgroups as well as on the copula C_m and $\widehat{C}_m$ to compute upper and lower bounds on the VaR of the sum $X_1 + \cdots X_{16}$, i.e. bounds on the distribution function $\mathbb{P}\Big(\sum_{i=1}^{16} X_i \le x\Big) = \mathbb{P}\Big(\sum_{j=1}^{m} X'_{I_j} \le x\Big)$. Therefore we use our reduction schemes in conjunction with the upper and lower copula bounds $C^{\theta_1}, C^{\theta_2}$ and $\widehat{C}^{\theta_1}, \widehat{C}^{\theta_2}$ as well as the marginals F'_j for $j = 1, ..., m$. The following table shows the improved standard bounds for different levels of the confidence threshold α along with the tightest bounds obtained by inverting $\underline{D}_\varphi(k), \underline{\widehat{D}}_\varphi(k)$ and $\overline{D}_\varphi(k), \underline{\widehat{D}}_\varphi(k)$, for $k = m + 1$ and $\varphi(x_1, ..., x_m) = \mathbb{1}_{x_1 + \cdots x_m < s}$, along the variable s. For the computation we assume that the parameters of the Frank copula amount to $\theta_1 = 0$ which is equivalent to the assumption of independence between the groups and $\theta_2 = 1$ which corresponds to weak positive dependence between the subgroups. For the sake of legibility, the results are rounded to integers.

	$m = 8$			$m = 4$			$m = 2$		
α	i. standard (low : up)	scheme (low : up)	impr. %	i. standard (low : up)	scheme (low : up)	impr. %	i. standard (low : up)	scheme (low : up)	impr. %
0.95	42 : 113	59 : 86	62%	22 : 150	39 : 112	43%	12 : 193	28 : 150	33%
0.99	82 : 210	108 : 147	70%	42 : 264	67 : 175	51%	21 : 329	42 : 218	43%
0.995	105 : 266	135 : 180	72%	53 : 329	83 : 206	55%	43 : 403	51 : 252	44%

Table 3: Improved standard bounds and VaR estimates via reduction schemes for $X_1 + \cdots + X_{16}$ given distributions of subgroups.

We identify a considerable improvement, ranging from 33% in the case of 8 groups consisting of $m = 2$ risks and $\alpha = 0.95$ to 72% for 2 groups with $m = 8$ risk factors. Note, that the improvement deteriorates as the number of subgroups increases. This implies that, with this configuration, a decline in the number of subgroups and the associated decrease in model uncertainty results in narrower risk bounds. Except for the case $\alpha = 0.95$ and $m = 2$, all of the bounds computed via our reduction scheme improve the unconstrained bounds in the marginals-only case which amount to 20–136 when $\alpha = 0.95$, 24–244 for $\alpha = 0.99$ and 25–303 for $\alpha = 0.995$. $\Diamond$

A. Using information on the survival copula

In this section we show that the reduction schemes in Section 4.1 and 4.2 can be applied similarly when information on the survival copula is provided. Specifically, we assume that the copula of $\mathbf{X}$ is such that $\widehat{\underline{Q}} \leq \widehat{C} \leq \widehat{\overline{Q}}$ where $\widehat{C}$ is the survival-function of C and $\widehat{\underline{Q}}, \widehat{\overline{Q}}$ are quasi-survival functions. We hence consider the generalized Fréchet functionals

$$\widehat{\underline{P}}_\varphi := \inf\left\{\mathbb{E}_C[\varphi]\colon C \in \mathcal{C}^d, \widehat{\underline{Q}} \leq \widehat{C} \leq \widehat{\overline{Q}}\right\}, \tag{A.1}$$

$$\widehat{\overline{P}}_\varphi := \sup\left\{\mathbb{E}_C[\varphi]\colon C \in \mathcal{C}^d, \widehat{\underline{Q}} \leq \widehat{C} \leq \widehat{\overline{Q}}\right\}. \tag{A.2}$$

Note that due to

$$\widehat{C}(F_1(u_1), ..., F_d(u_d)) = \mathbb{P}(X_1 > u_1, ..., X_d > u_d) = \mathbb{P}(-X_1 < -u_1, ..., -X_d < -u_d),$$

for all $\mathbf{u} \in \mathbb{I}^d$, the condition $\widehat{\underline{Q}} \leq \widehat{C} \leq \widehat{\overline{Q}}$ is equivalent to $\underline{Q} \leq C_{-\mathbf{X}} \leq \overline{Q}$, where $\underline{Q}(\mathbf{u}) := \widehat{\underline{Q}}(\mathbf{1} - \mathbf{u})$, $\overline{Q}(\mathbf{u}) := \widehat{\overline{Q}}(\mathbf{1} - \mathbf{u})$ and $C_{-\mathbf{X}}$ is the copula of $-\mathbf{X}$. In particular, since $\underline{Q}$ and $\overline{Q}$ are quasi-copulas it follows from our duality theorem 3.2 and a transformation of variables that the sharp dual bound corresponding to $\widehat{\underline{P}}_\varphi$ is given by

$$\widehat{\underline{P}}_\varphi = \widehat{\underline{D}}_\varphi = \sup\left\{\widehat{\underline{Q}}(h) - \widehat{\overline{Q}}(g^-) + \sum_{i=1}^d \mathbb{E}_i[f_i]\colon (h, g, f_1, ..., f_d) \in \widehat{\underline{\mathcal{A}}}\right\}, \tag{A.3}$$

where

$$\widehat{\underline{\mathcal{A}}} := \left\{(h, g, f_1, ..., f_d)\colon f_i \in L(F_i),\ i = 1, ..., d,\ h, g \in \widehat{\mathcal{R}} \text{ s.t. } h - g^- + \sum_{i=1}^d f_i \leq \varphi\right\}.$$

and

$$\widehat{\mathcal{R}} := \left\{h = \sum_{n=1}^k \alpha_n \widehat{\Lambda}_{\mathbf{u}^n}\colon k \in \mathbb{N}, \alpha_1, ..., \alpha_k \geq 0, \mathbf{u}^1, ..., \mathbf{u}^k \in \overline{\mathbb{R}}^d\right\},$$

for $\widehat{\Lambda}_{\mathbf{u}}$ of the form

$$\widehat{\Lambda}_{\mathbf{u}}\colon \mathbb{R}^d \ni (x_1, ..., x_d) \mapsto \mathbb{1}_{x_1 \geq u_1, ..., x_d \geq u_d}.$$

Moreover, we denote $\widehat{\Lambda}^-_{\mathbf{u}}(x_1, ..., x_d) := \mathbb{1}_{x_1 > u_1, ..., x_d > u_d}$ and h^- for $h \in \widehat{\mathcal{R}}$ is defined accordingly. Finally, for quasi-survival functions $\widehat{Q}$ and $h = \sum_{n=1}^k \alpha_n \widehat{\Lambda}_{\mathbf{u}^n} \in \mathcal{R} \in \widehat{\mathcal{R}}$ we define,

$$Q(h) := \sum_{n=1}^k \alpha_n \widehat{Q}\big(F_1(u_1^n), ..., F_d(u_d^n)\big); \quad Q(h^-) := \sum_{n=1}^k \alpha_n \widehat{Q}\big(F_1^-(u_1^n), ..., F_d^-(u_d^n)\big).$$

Analogously, the sharp dual bound associated to $\widehat{\overline{P}}_\varphi$ is equal to

$$\widehat{\overline{P}}_\varphi = \widehat{\overline{D}}_\varphi = \inf\left\{\widehat{\overline{Q}}(h^-) - \widehat{\underline{Q}}(g) + \sum_{i=1}^d \mathbb{E}_i[f_i]\colon (h, g, f_1, ..., f_d) \in \widehat{\overline{\mathcal{A}}}\right\}, \tag{A.4}$$

where

$$\widehat{\widehat{\mathcal{A}}} := \Big\{(h,g,f_1,...,f_d)\colon f_i \in L(F_i),\ i=1,...,d,\ h,g \in \widehat{\mathcal{R}} \text{ s.t. } h^- - g + \sum_{i=1}^d f_i \geq \varphi\Big\}.$$

Based on these dual characterizations the following corollaries establish the corresponding reduction schemes. Using the fact that

$$\mathbb{P}(\psi(X_1,...,X_d) < s) = 1 - \mathbb{P}(\psi(X_1,...,X_d) \geq s),$$

the proofs involve similar arguments as the proofs of Theorem 4.6 and Theorem 4.10 and therefore they are omitted. We denote the componentwise maximum of vectors $\mathbf{u}^1,...,\mathbf{u}^k \in \mathbb{R}^d$ by

$$\max(\mathbf{u}^1,...,\mathbf{u}^k) = \Big(\max_{n=1,...,k}\{u_1^n\},..., \max_{n=1,...,k}\{u_d^n\}\Big).$$

Corollary A.1. *Let $\varphi(x_1,...,x_d) = \mathbb{1}_{\psi(x_1,...,x_d)<s}$ for $\psi\colon \mathbb{R}^d \to \mathbb{R}$ increasing in each coordinate and let*

$$\begin{aligned}\widehat{\widehat{D}}_\varphi(k) := \inf\Big\{1 - &\sum_{\mathbf{u}\in\widehat{\mathcal{M}}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+\, \underline{\widehat{Q}}\big(F(\mathbf{u})\big) \\ &- \sum_{\mathbf{u}\in\widehat{\mathcal{M}}^e(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+\, \widehat{\overline{Q}}\big(F(\mathbf{u})\big)\colon \mathbf{u}^1,...,\mathbf{u}^k \in \mathcal{U}_\psi^c(s)\Big\},\end{aligned} \tag{A.5}$$

where

$$\begin{aligned}\widehat{\mathcal{M}}^o(\mathbf{u}^1,...,\mathbf{u}^k) &:= \{\max(\mathbf{u}^{i_1},...,\mathbf{u}^{i_m})\colon 1 \leq i_1 < \cdots < i_m \leq k\ \ m \text{ odd}\},\\ \widehat{\mathcal{M}}^e(\mathbf{u}^1,...,\mathbf{u}^k) &:= \{\max(\mathbf{u}^{i_1},...,\mathbf{u}^{i_m})\colon 1 \leq i_1 < \cdots < i_m \leq k\ \ m \text{ even}\}.\end{aligned}$$

and

$$\begin{aligned}l^o(\mathbf{u}) &:= |\{(i_1,...,i_m)\colon 0 \leq i_1 < \cdots < i_m \leq k,\ m \text{ odd}, \mathbf{u} = \max(\mathbf{u}^{i_1},...,\mathbf{u}^{i_m})\}|,\\ l^e(\mathbf{u}) &:= |\{(i_1,...,i_m)\colon 0 \leq i_1 < \cdots < i_m \leq k,\ m \text{ odd}, \mathbf{u} = \max(\mathbf{u}^{i_1},...,\mathbf{u}^{i_m})\}|,\end{aligned} \tag{A.6}$$

for $\mathbf{u} \in \widehat{\mathcal{M}}^o(\mathbf{u}^1,...,\mathbf{u}^k)$ and $\mathbf{u} \in \widehat{\mathcal{M}}^e(\mathbf{u}^1,...,\mathbf{u}^k)$ respectively. Then it holds that

$$\widehat{\widehat{D}}_\varphi(k) \geq \widehat{\widehat{D}}_\varphi(k+1) \geq ... \geq \widehat{\widehat{D}}_\varphi.$$

Remark A.2. Corollary A.1 extends the upper improved standard in the sense that

$$\inf_{h\in\widehat{\mathcal{R}}} 1 - \underline{\widehat{Q}}(h) = \inf_{\mathbf{u}\in\mathcal{U}_\psi^c(s)} 1 - \underline{\widehat{Q}}(F(\mathbf{u})) = \overline{M}_{\underline{\widehat{Q}},\psi}(s).$$

The following corollary establishes a similar reduction scheme for $\underline{\widehat{D}}$. To this end, let us denote for $\mathbf{u}^1,...,\mathbf{u}^k \in \mathbb{R}^d$ and an enumeration $\{\mathbf{i}^1,...,\mathbf{i}^{dk}\}$ of $\{1,...,d\}^k$

$$\widehat{\mathrm{U}}_{\mathbf{i}^n} := \max(\widehat{\mathrm{pr}}_{i_1}(\mathbf{u}^1),...,\widehat{\mathrm{pr}}_{i_k}(\mathbf{u}^k)) \quad \text{for } (i_1,...,i_k) = \mathbf{i}^n, n = 1,...,dk,$$

where $\widehat{\mathrm{pr}}_i(\mathbf{u}) := (-\infty,...,-\infty,u_i,-\infty,...,-\infty)$ for $i \in \{1,...,d\}$.

Corollary A.3. *Let* $\varphi(x_1,...,x_d) = \mathbb{1}_{\psi(x_1,...,x_d)<s}$ *for* $\psi\colon \mathbb{R}^d \to \mathbb{R}$ *increasing in each coordinate and let*

$$\begin{aligned}\underline{\widehat{D}}_\varphi(k) := \sup\Bigg\{ 1 - &\sum_{\mathbf{u}\in\widehat{\mathcal{W}}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \,\widehat{\overline{Q}}\big(F(\mathbf{u})\big) \\ &- \sum_{\mathbf{u}\in\widehat{\mathcal{W}}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \,\underline{\widehat{Q}}\big(F(\mathbf{u})\big) \colon \mathbf{u}^1,...,\mathbf{u}^k \in \mathcal{U}_\psi(s) \Bigg\},\end{aligned} \tag{A.7}$$

where

$$\begin{aligned}\widehat{\mathcal{W}}^o(\mathbf{u}^1,...,\mathbf{u}^k) &:= \widehat{\mathcal{M}}^o(\widehat{\mathrm{U}}_{\mathbf{i}^1},...,\widehat{\mathrm{U}}_{\mathbf{i}^{dk}}),\\ \widehat{\mathcal{W}}^e(\mathbf{u}^1,...,\mathbf{u}^k) &:= \widehat{\mathcal{M}}^e(\widehat{\mathrm{U}}_{\mathbf{i}^1},...,\widehat{\mathrm{U}}_{\mathbf{i}^{dk}}),\end{aligned} \tag{A.8}$$

for an enumeration $\{\mathbf{i}^1,...,\mathbf{i}^{dk}\}$ *of* $\{1,...,d\}^k$ *and* l^o, l^e *are given in* (A.6)*. Then it holds that*

$$\underline{\widehat{D}}_\varphi(k) \geq \underline{\widehat{D}}_\varphi(k+1) \geq ... \geq \underline{\widehat{D}}_\varphi.$$

References

D. Bartl, P. Cheredito, M. Kupper, and L. Tangpi. Duality for increasing convex functionals with countably many marginal constraints. Preprint, arXiv:1509.08988v2, 2015.

D. Bartl, M. Kupper, T. Lux, and A. Papapantoleon. Sharpness of improved Fréchet-Hoeffding bounds: an optimal transport approach. Preprint, arXiv:1709.00641, 2017.

M. Beiglböck, P. Henry-Labordère, and F. Penkner. Model-independent bounds for option prices: A mass transport approach. *Finance Stoch.*, 17:477–501, 2013.

C. Bernard and S. Vanduffel. A new approach to assessing model risk in high dimensions. *J. Banking Finance*, 58:166–178, 2015.

V. Bignozzi, G. Puccetti, and L. Rüschendorf. Reducing model risk using positive and negative dependence assumptions. *Insurance Math. Econ.*, 61:17–26, 2015.

Board of Governors of the Federal Reserve System. Supervisory guidance on model risk management. technical report, URL: www.federalreserve.gov/supervisionreg/srletters/sr1107.htm, 2011.

M. Cambou and D. Filipović. Model uncertainty and scenario aggregation. *Math. Finance*, 2015. forthcoming, doi:10.1111/mafi.12097.

M. Denuit, C. Genest, and É. Marceau. Stochastic bounds on sums of dependent risks. *Insurance Math. Econom.*, 25:85–104, 1999.

P. Embrechts and M. Hofert. A note on generalized inverses. *Math. Methods Oper. Res.*, 77:423–432, 2013.

P. Embrechts and G. Puccetti. Bounds for functions of dependent risks. *Finance Stoch.*, 10:341–352, 2006.

P. Embrechts, A. Höing, and A. Juri. Using copulae to bound the value-at-risk for functions of dependent risks. *Finance Stoch.*, 7:145–167, 2003.

P. Embrechts, G. Puccetti, and L. Rüschendorf. Model uncertainty and VaR aggregation. *J. Banking Finance*, 37:2750–2764, 2013.

N. Gaffke and L. Rüschendorf. On a class of extremal problems in statistics. *Math. Operationsforschung Statist.*, 12:123–135, 1981.

S. Das Gupta, M. L. Eaton, I. Olkin, M. Perlman, L. J. Savage, and M. Sobel. Inequalitites on the probability content of convex regions for elliptically contoured distributions. In *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, pages 241–265. Univ. of Calif. Press, 1972.

H. G. Kellerer. Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete*, 67:399–432, 1984.

J. A. De Loera, R. Hemmecke, and M. Köppe. *Algebraic and Geometric Ideas in the Theory of Discrete Optimization*. MOS-SIAM, 2013.

T. Lux and A. Papapantoleon. Improved Fréchet–Hoeffding bounds on d-copulas and applications in model-free finance. Preprint, arXiv:1602.08894v3, 2015.

T. Lux and A. Papapantoleon. Model-free bounds on value-at-risk using partial dependence information. Preprint, arXiv:1610.09734, 2016.

R. B. Nelsen. *An Introduction to Copulas*. Springer, 2nd edition, 2006.

R. B. Nelsen, J. J. Quesada-Molina, J. A. Rodriguez-Lallena, and M. Ubeda-Flores. Bounds on bivariate distribution functions with given margins and measures of associations. *Comm. Stat. Theory Methods*, 30:1155–1162, 2001.

G. Puccetti and L. Rüschendorf. Computation of sharp bounds on the distribution of a function of dependent risks. *J. Comput. Appl. Math.*, 236:1833–1840, 2012a.

G. Puccetti and L. Rüschendorf. Bounds for joint portfolios of dependent risks. *Statistics and Risk Modeling*, 29:107–132, 2012b.

G. Puccetti and L. Rüschendorf. Sharp bounds for sums of dependent risks. *J. Appl. Probab.*, 50:42–53, 2013.

G. Puccetti, L. Rüschendorf, and D. Manko. VaR bounds for joint portfolios with dependence constraints. *Dependence Modeling*, 4:368–381, 2016.

G. Puccetti, L. Rüschendorf, D Small, and S. Vanduffel. Reduction of value-at-risk bounds via independence and variance information. *Scand. Actuar. J.*, 2017:245–266, 2017.

S. T. Rachev and L. Rüschendorf. Solution of some transportation problems with relaxed or additional constraints. *SIAM J. Control Optim.*, 32:673–689, 1994.

S. T. Rachev and L. Rüschendorf. *Mass Transportation Problems, Volume I: Theory*. Springer, New York, 1998.

L. Rüschendorf. Sharpness of Fréchet-bounds. *Z. Wahrsch. Verw. Gebiete*, 57:293–302, 1981.

L. Rüschendorf. Monge-Kantorovich transportation problem and optimal couplings. *Jahresbericht der DMV*, 109:113–137, 2007.

L. Rüschendorf and J. Witting. VaR bounds in models with partial dependence information on subgroups. *Dependence Modeling*, 5:59–74, 2017.

A. Syropoulos. Mathematics of multisets. *Lecture Notes in Computer Science*, 2235:347–358, 2001.

P. Tankov. Improved Fréchet bounds and model-free pricing of multi-asset options. *J. Appl. Probab.*, 48: 389–403, 2011.

C. Villani. *Optimal Transport. Old and New*. Springer, Heidelberg, 2009.

B. Wang and R. Wang. The complete mixability and convex minimization problems with monotone marginal densities. *J. Multivariate Anal.*, 102:1344–1360, 2011.

R. C. Williamson and T. Downs. Numerical methods for calculating convolutions and dependency bounds. *Internat. J. Approx. Reason.*, 4:89–158, 1990.